考试学

THE SCIENCE OF TESTING

廖平胜　何雄智　梁其健

华中师范大学出版社

内容提要

《考试学》是一部首次系统探究考试科学理论体系的学术专著。它以中外古今丰厚的考试材料为依据，紧密结合当今人类考试实践面临的崭新课题，从历史与现实、理论与实践的结合上，认真探讨了人类考试产生发展的历史轨迹及演变成因，明确阐述了考试的内涵、要素、属性、本质与功能，具体分析了教、考、学三者的内在联系及社会人才系统育才、选才、用才三大环节的运行机制，以及考试与考生、教育、人事管理、人才、家庭和社会环境的关系，系统介绍了考试机构的类别、人员结构与职能，考试的设计、实施与质量分析，各类考试的特点、程序、原则、方法、技术手段与适用范围，并提出了学校考试和十类人事考试的模式。

本书的一个显著特点，是从科学的角度探讨人类考试活动的客观规律，致力寻求现行考试有效实践的理论依据和科学途径，注重实用。可供各级各类考试、人事、教育、教学、行政等方面的管理工作者和研究工作者参考、借鉴，也可用作高校相关专业的教材。

SYNOPSIS

The Science of Testing is the first attempt to study the system of the theory of testing, to combine history with reality and to combine theory with practice.

Based on the rich data at different times and in different countries, it traces the origin and development of testing, examines the changes and their causes. It clearly defines the content, factors, nature and functions of testing. In this book, careful analysis has been made of the inner relationship among teaching, testing and learning, of the operation devices in training, selecting and using talented persons, and of the relationship among testing, testees, education, management of human beings, families and social situations. It describes the types of testing organizations, the structure and duties of testers, the construction and operation of testing, and the conducting of score analysis. It also discusses the features, precedur-

es, principles, methods, technical devices and scopes of different kinds of testing. The authors have proposed models for educational testing and other kinds of testing.

One of the prominent features of the book is trying to find out the objective laws, the theoretical basis, and scientific approaches of the present testing practice from the point of view of science while paying attention to practical uses. So this book can be used as a reference book by testers, managers, educationists, teachers, administrative and research workers. It also can be used as a textbook for students who will be involved in testing in the future.

前　言

回首人类社会演进的历程，迄今未有全无考试活动的社会，离开人类社会的考试活动也从来没有。考试这种历来为世人关注而又颇富魅力的社会活动，它与人及其所组成的具有交互和共同关系的社会密不可分。自从有了人——会思维且能有意识从事社会实践活动的人，就有了鉴别人的德、学、才、识、体诸方面差异的考试，人类一经产生文字，也就出现了有关考试史实与认识的记述，后随人类社会的兴衰起伏而历经沧桑，如今“考试已是这样稳固地站定了脚跟，要废除它，似乎比取消篝火节或者圣诞节更无可能”，它已渗透于社会生产与社会生活各个领域的不同阶层，成为人们有效开发和科学利用社会人力资源、促进社会发展、迎接富有挑战性的未来的重要手段。

数千年来，考试尤其是各行各业的选拔与合格考试，不仅是人们争相角逐的目标、希望视线的焦点，而且始终是人类科学研究中具有特殊诱惑力的课题。为强化考试的社会适应能力，使之正确发挥不同社会期求之功能，各个时代的人们都曾倾心竭智于此。但是，由于人类科学技术及考试自身发展的限制，历史上有关考试认识的阈限，未曾越出考试历史、内容、方法、手段的疆界，传给后世的大都是经过某种修补的祖宗之法，而缺乏科学考试理论的指导，故在新的考试实践中仍旧弊病重犯、失误甚多。

时至本世纪70年代末期，现代科学为人类社会带来昌明的事实，以及现代社会赋予考试研究以新的历史使命的客观要求，导致人们对考试认识及研究宗旨的深刻变革，认为考试并非是简单的工具或手段，也远不止是一个方法或技术问题，而是人类社会特有的现象，一种主试与被试两种主体彼此交互的社会活动，

有其固有的客观规律。因此，深感对考试史实的陈述、经验的总结与介绍、方法和手段的探讨与传递，仅是探讨考试革新之进程的举步台阶，而不是考试研究的终极目的。其终极目的，是通过纷繁的考试现象揭示考试的本质，从考试演变的史实中接受启示，并对其合理的思想加以承继与发挥，进而形成反映考试规律的科学理论体系，以推动考试实践的发展。于是，人们把革新考试的理想由方法、技术的探索转托于科学理论的研究，使考试学迅速成为中外理论界潜心研讨的热门科学。

由于上述思想的感染，以及我国改革传统考试之呼声的鼓动，我们于1984年在国内首倡并着手考试学的创建。几年来，在国家教委学生司、劳动人事部干部局、中央教科所、湖北、陕西、四川、黑龙江、吉林、上海、辽宁、山西、重庆等省市高校招生办公室和自学考试办公室的大力支持下，在华中师范大学科研处、教务处、招生办公室、教科所的领导和资助下，我们写出了我国研究考试科学的第一部理论专著——《考试学》。

考试学作为一门科学，它的本质任务是揭示考试的客观规律，以这些规律的知识指导新的考试实践则是其根本目的。在《考试学》一书中，我们本着源于实践、受检于实践、服务于实践的原则，就什么是考试，考试产生、发展的演变成因，考试与人的智能发展、社会经济、产业、技术、职业结构的关系，考试与教育、人事管理的关系，考试观、人才观、教育观的演变及其关联，考试的要素与属性，考试的类别与方法，考试设计与实施的程序，考试数据整理与质量分析，各类考试的原则与方法，考试机构的类别与职能，考试机构的设置、人员的结构与素质，社会各系统人事考试的模式、类别与原则，考试在人类社会发展中的功能等问题，分别进行了理论阐述和实际应用的探讨。

但是，勿庸讳言，书中所述仅是我们自己的一些浅见。由于历史上没有同类书籍可资借鉴，现今有关考试科学理论的探讨，

无论是研究方法还是研究成果，都还处在襁褓时代，要开拓考试科学理论体系这一新领域，其进程的每一步都充满了困难与艰辛，对此，我们在撰写这部著作的开笔之时，就已充分认识到。然而，改变我国考试理论落后状态、革新传统考试模式的时代要求，以及人们为创立具有中国特色的现代考试系统而奋力求索的精神，又鞭策我们为中国考试科学理论体系的形成与发展而贡献出自己的绵薄之力。这就是我们不揣谫陋，为创建考试学而撰写这部理论著作的动机。“初生之物，其形必丑。”但由丑至美亦是事物发展的一般规律。这部《考试学》问世后，我们恳请同行专家和广大读者批评、指正，期盼早日出现更加完美的考试理论著作，以满足社会发展的需求。

承担本书撰写任务的是：廖平胜第一、三、六、七、八、九、十章，何雄智第二、五章，梁其健第四、十一章。全书由廖平胜副教授统稿。

在撰写及修改本书的过程中，陶军教授和杨葆焜教授给予了具体指导，得到了不少专家学者的鼓励和帮助，潘香春同志参与了资料收集工作，日本大学入学考试国家中心副委员长肥田野直教授和香港考试局副秘书陈永照先生给予了资料支持，同时参阅了中外有关书刊资料。在此，谨向上述所有同志致以最诚挚的谢意。

作　者
1987年4月于武汉

目　录

第一章

总　　论

一门学科的产生与发展，有赖于人类生产和生活客观需求这一巨大原动力的推动；一门学科的兴衰，决定于它能够满足社会需求的程度；而要使一门学科从经验的总结上升为科学的理论体系，则取决于研究的方法和手段。

考试是人类社会特有的现象，是测度、甄别人的德、学、才、识、体个别差异的一种社会活动。它产生于社会生产和社会生活的客观需要，并随社会发展需求的变化而不断改变其内容、性质和模式，形成与其社会相适应的结构，以发挥不同社会或时代所期求的功能。考试和教育是一对孪生子，在人类社会的母系统中，它们始终是彼此独立而又密切关联的两个子系统，始终是开发智力资源和人力资源的基本手段，始终是社会进化的重要机制。

自古以来，没有全无考试活动的社会，也没有从未进行过考试的民族或国家。考试这种人类社会的特有现象，既是一个普遍的范畴，又是一个永恒的范畴。凡人们集体从事物质和精神生产的地方，就必有考试现象存在，只要人类社会传递发展文化、判定人的德、学、才、识、体个别差异的客观需求不终止，考试

现象也就不会消亡。

考试是社会的产物，为一切社会所必需。在人类社会发展的历史长河中，虽有社会形态的更替，但任何一种社会形态都“是人们交互作用的产物”①。生活在不同社会或时代的人们，其物质生活的生产方式虽各有所异，然而，任何时代的人们为了社会生命的继续不绝，都有一个改变现状、创造未来的问题，也无不面临着同样的社会挑战：如何使人类文化绵延不断？如何使新生一代获得与其社会要求相适应的文化知识和技能，并承担起社会所赋予的历史使命？如何使不同素质与智能的人各尽其才、才尽其用？解决这些问题的途径，无不仰赖振兴教育和革新考试。我们不能设想，在一个没有教育的社会，人类的文化还能得以保持、传递和发展，新生一代可自我造化成符合社会发展要求的人才，能负起推动社会前进的使命。同样，我们也不能设想，在一个缺少育才与用才中介——考试的社会，教育能够圆满实现预期目标，社会人力资源可得到合理的开发和利用。正如日本著名学者梶田睿一在《现代教育评价论》一书中所指出，倘若把以考试为核心的所有评价措施都清除，“就将出现这种情况：尽管连读、写、算这样的基本技能都没掌握，然而还自以为‘得计’，苟且偷安、自我满足、而又怠惰不逊的年轻人，将被大批地输送到社会上来，……这样下去，其结果将导致整个社会的崩溃！”考试在人类社会发展中的作用并非人为，而是社会生产和社会生活的客观需求，也是考试自身的固有功能。科学的考试，不仅是教学双边活动联系的纽带，督导教学活动、检验教学效果、实现教育目标的有效措施，而且是人类社会辨别人才、发现人才、评价人才、选拔人才必不可少的工具，社会各系统人事管理部门获取决定各类工作人员培养、录用、晋升、革职、调迁、授衔、奖惩等客

①《马克思恩格斯选集》第4卷，人民出版社1972年版，第320页。

观依据的基本途径，调节社会各部门人才结构的有力杠杆，开发利用社会人力资源、提高经济效益的重要手段。因而，考试历来为世人所关注。

第一节 考试是一门科学

对考试这一人类社会的特有现象，古今中外不乏探究之人。长期以来，人们在教育学、心理学、人才学、社会学、管理学、行政学等学科的研究中，都曾涉及到考试问题。前人的许多历史论著中，也有大量社会及学校考试史实的记述。至于考试制度和方法，更是为历代考试研究工作者所勤苦耕耘。但是，无论是考试方法、考试制度历史及现实的研究，还是教育学、心理学、社会学、人才学、管理学、行政学等学科对考试问题的论述，都是把考试作为一种工具、手段或方法探讨的，对于考试的内在规律至今尚未被揭示。现阶段，违背考试规律、不符考试原则的现象仍较为普遍，一些陈腐的考试观念和不科学的考试程序、方法与手段，已成为现实育才、选才和用才的严重障碍。尽管"恶性补习"、"升学至上"、"高分低能"、"智能偏枯"、"文凭贬值"、"能级不称"种种弊端，是多种社会因素所致，但缺少科学的考试理论、方法和手段，不能不是根本原因之一。因此，如何正确发挥考试的教育性，为亿万后继人才智能的全面发展导以正确方向？如何提高考试的科学性，使人才的甄选准确、任使恰当？如何有效发挥考试的调节作用，使社会各部门的人才结构合理化，不断提高人力资源的利用率和人才使用的经济效益？如何揭示考试的内在规律，实现考试本质因素的外化，借以指导考试的改革，尽快建立符合现代社会发展要求的考试模式？诸如此类的问题，仍是当今考试研究所面临着的严峻挑战，这就要求我们

从马克思主义的立场出发，创造性地而不是教条式地去寻求能为现实社会采纳的回答。同时也说明，考试的研究必须进入理论的领域，“在这里只有理论思维才能有所帮助”①，才能担负起揭示考试现象客观规律的任务。由此可见，将考试作为一门独立的科学进行研究，既是社会发展的客观需要，又是考试自身发展的必然，同时也是现代科学高度综合而又不断分化的结果。

人类科学发展的历史表明，在其发展的不同历史阶段，都有不同的特点。当人类社会进入20世纪60年代后，由于现代科学迅猛发展，新发现、新发明、新成果与日俱增，其发展态势呈现出互相对立而又相互联系的两大特点：一是各门学科之间不断综合，进而形成具有新的研究对象和更大研究范围的综合性学科，为学科的重新分化创设条件；另一方面，各门学科本身不断分化，产生具有特殊对象、内容和方法并自成体系的新兴学科，为学科的进一步高度综合奠定基础。这两者互为条件、相互促进、互为依存又彼此对立的矛盾运动，便构成了现代科学发展的一种推动力，也是考试学问世的得力“助产士”。长期以来，由于考试自身发展和人们对考试现象认识的局限，一直把考试仅看作一种手段、方法或工具，将其置于教育学、心理学、社会学、人才学、管理学、行政学等学科的研究范围之中，或被视为这些学科的一个邻接领域。随着科学技术的发展，现代物质生产不断赋予考试以新的内容和新的使命，使考试在更为广阔的社会领域得到发展，其体系更加完善，人们对考试的研究和认识也随之深化，愈来愈感到，教育学、心理学、社会学、人才学、管理学、

①《马克思恩格斯选集》第3卷，人民出版社1972年版，第465页。

行政学等有关考试的阐述，旨在为其所研究的对象服务，它们没有也不可能揭示考试现象的本质和规律，满足不了现代社会对考试科学理论的需求。考试因此打破了原居学科的“壁垒”，只身走出了教育等学科的领域，开始了独立成长的生涯。正象恩格斯当年在推断科学发展方向时所预言的那样：“原有学科的邻接领域将是新学科的生长点”①，促使人们在这学科的邻接处进行新的探索和开拓，“如地球物理、生物物理、生物化学、天体物理、宇宙生物学、生物地球化学等等”②，就是在数学、物理学、化学、天文学、地理学、生物学这些基础学科的邻接领域生长起来的。不仅如此，近几十年来，在一些分支学科之间，乃至在大部类学科之间，同样有许多边缘学科产生。因此，为目前许多国家潜心研究的热门学科——考试学之所以能在现代科学高度综合而又不断分化的过程中产生，并为世界上日益众多的人们所关注和承认，也就不足为奇了。

深入开展考试科学的研究，其意义远不只是考试本身，离开了系统的科学理论指导，就不可能有科学的实践。系统科学的考试理论，对于教学内容与方法的改革，课程结构的更新，教育评估系统的建立，考试制度及方法的革新与完善，人事管理手段的改进，社会各部门人才结构的调整，民族文化素质的提高等，都有积极的促进作用。考试学研究的最终目标，是建立考试科学的理论体系，用以指导现代社会的考试实践，使考试更好地服务于社会生产与社会生活。

一、考试学的对象和任务

考试学是一门研究考试现象及其规律的新兴社会科学。它以

①②《新技术革命与管理》，湖北科技出版社1984年版，第12页。

人类测度、甄别人的德、学、才、识、体个别差异的社会活动——考试活动为研究对象。探求考试发生、发展的演变成因，研究考试的内在结构、外部关系及功能，揭示考试的客观规律，论述考试的基本理论、原则、程序与方法，以指导考试实践，并为考试科学的各分支学科奠定理论基础，乃是考试学的主要任务。

"考试"一词有广、狭义之分。广义的考试，泛指人类社会一切测度、甄别人的知识、技能、能力、智力、品质、体质、体能等群体或个别差异的活动。狭义的考试，即是由主试者根据一定社会的要求，在一定的场所，采取一定的方式方法，选择适当的内容，对应试者的德、学、才、识、体（诸方面或某方面）所进行的有组织、有目的的测度或甄别活动。

人类考试形式众多、种类极繁，大致可归为两大系统：社会考试系统和学校考试系统。这两大系统所含各种类别和形式的考试，均属考试学讨论的范围。

二、考试学的基本体系、内容和性质

（一）考试学的基本体系和内容

考试作为一门科学，它是许多国家近五、六年来才致力开拓的一个新领域。在国外，由于新技术革命的蓬勃兴起，旧的物质生产方式受到剧烈冲击，从而导致技术结构、产业结构、企业结构、职业结构以及从业人员素质与智能结构的深刻变革。迅速提高劳动者的素质和智能，准确测度劳动者的质量，恰当选任符合现代生产要求的人才，日益成为决定生产发展进程的关键。因此，在本世纪70年代末期，一些发达的资本主义国家便把考试视为迎接新技术革命挑战战略决策研究中的重大理论课题。现阶段，考试学已是国外苦心探索的一门新兴科学，其理论体系正在形成之中，只是尚无系统的理论专著问世。在我国，从科学的角度对考试进行系统研究始于80年代初。由于国家和各级政府的

积极倡导，以及社会各界的大力支持，考试研究很快推向全国。特别是近年来经济、科技、教育、政治体制改革的深入展开，加快了我国现代化建设的步伐，经济、技术、产业、职业结构随之改变，社会对各类劳动者素质与智能的要求迅速提高，出现了历史上罕见的热切渴望人才的局面，从前单凭体力拼搏的生产观念日趋消退，以知识、智力去创造物质和精神文明的思想代之而起，一种新的人才观由此而生。现今各行各业对人才的培养和选用，无不注重质量及人力投资的经济效益，无不强调用人的客观标准，无不通过考试去测度或甄别人才的知识、技能、能力、品质与体能，以求人才培养合格、甄选准确、任使恰当、管理得法。现实社会的客观需求，促使考试如雨后春笋般发展起来，以至渗透于社会生产与生活的各个领域。

一门科学的发展总是取决于社会需求的程度。考试在我国的空前繁荣，不仅是对传统考试的全面检测，使其利弊明显暴露于外，而且成了我国考试科学研究的催化剂，使考试研究由单一化、分散化走向综合化、系统化，从经验总结跃入理论研讨的大门，学术论文大量见诸于报刊，且有少量文集印行，一些孕育中的考试论著渐趋成熟，考试科学的理论体系正在探索创建中。《考试学》一书首次斗胆提出一个初步模式，旨在抛砖引玉，以期中国考试科学完整理论体系的早日成熟。

考试学的内容涉及社会、自然及应用技术科学领域的多门学科，但社会科学的成分是其主体。更为重要的是，我国的考试学，是我国社会主义物质文明和精神文明建设时期的产物，它必以辩证唯物主义的哲学理论为指导，从中国的国情出发，在合理继承传统考试理论科学成分、认真总结新中国考试实践经验、择善吸收外国考试理论研究新成果的基础上，系统深入地探讨和揭示考试的内在规律，建立富有中国特色的考试科学理论体系，为创建符合中国社会发展需要的现代学校考试系统和现代人事考试系统服

务。本着这一原则，我们从纵横两个角度构思并形成了本《考试学》的基本体系：

纵的方面，根据考试和社会发展的规律，阐述考试的发生、发展与社会发展的关系，以及考试自身发展的历史轨迹和演变成因。

横的方面，通过讨论考试与人才、考生、教育、人事管理诸方面的关系，考试的基本体系、类别、原则、方法和技术手段，考试的组织机构及其职能，考试的教学功能与社会功能等，探讨考试的内、外部规律，阐述考试在人类社会中的实际应用。

在结构上，《考试学》分为三个层次：首先论述科学的考试对实现教育目标、提高民族素质、调节人才及智能结构、开发利用人力资源、促进人类社会两个文明建设的作用与影响。其次，论述考试的基本理论，如考试的本质、要素、属性、原则、类别、形式、程序、方法、组织机构等。最后，论述考试在人才培养、甄选、任使、管理中的具体使用。

在内容上，《考试学》包括四大方面。相对社会科学领域的其它学科来说，考试学是一门独立的学科，但相对考试科学中的分支学科而言，它又是一门综合性的基础理论学科。考试学所要研究的是考试这一社会实践活动的共性理论问题，揭示和反映那些带有普遍意义的一般规律。对于人类考试活动的悠久历史、古今中外至为繁复的考试制度、社会各行各业类型众多的具体考试等，它不可能作详尽的叙述和细致的讨论，拟在已经着手研讨的中国考试史、外国考试史、学校考试学、人事考试学等课题中另作阐述。因此，本书除在第一章对考试学的对象、任务、性质、原则、方法、依据以及考试学与临近学科的关系进行概述外，主要论述四大方面的内容：考试的教学功能和社会功能；考试学的基本原理；考试学的基本体系；考试的实际应用。

（二）考试学的性质

一门学科的性质为它研究的对象所决定。要弄清考试学的性质，首先必须明确考试现象本身的性质。

考试这种人类社会的特有现象，它作为一种测度、甄别人的德、学、才、识、体个别差异的社会活动，自古有之，不仅有其发生、发展的内在规律，而且有它固有的属性。就其社会职能讲，它既为社会生产服务，又为社会生活服务，（在阶级社会中为阶级斗争服务）；从其发展的基本规律看，它既受社会发展规律制约，同时又受人的发展规律制约。而考试作为一门科学，其性质因考试的社会属性和自然属性所决定，它属社会科学的性质，同时又兼有一定的自然科学的性质；既有鲜明的阶级性，又有强烈的实践性。

世间一切事物都是一个有规律的发展过程，考试也不例外，它同样是作为过程而向前延伸的。考试与教育同时问世，它们在生产力与生产关系的矛盾运动中不断获得发展，伴随人类社会的始终。因之，作为社会现象的考试，它总是受制于一定社会生产力和科学文化的发展水平，有什么样的经济结构、政治制度和科学文化水平，就有什么样性质的考试。

在原始社会早期，由于生产力水平极为低下，"生存的困难，同自然斗争的困难使原始人受到十分沉重的压抑"，①他们只能凭着群体的智慧和力量维持生存，过着"共寒共寒，共饥其饥"②的生活，没有剩余产品，也无阶级之分。劳动技术与生活经验十分贫乏。这些经验和技术的传授只需也只能在生活和生产的过程中进行，对于新一代社会成员接受经验快慢、掌握生产技术熟练程度、生活品质优劣的检验，尚处于直觉观察的自然形态，不含有阶级的意识倾向。这种萌芽状态的考试现象，也无所谓阶级性。

到了原始社会末期，由于生产工具的改进和劳动技能的提

①《列宁全集》第5卷，人民出版社1959年版，第89页。

②《尉缭子·治本》。

高，社会生产力有了新的发展，从而引起了人类第一次大分工，产品出现剩余，共劳共享的平等生活为贫富所分化，统治与被统治两大阶级的对立日渐明显，考试的性质便因此发生变化，即由一种无阶级意识的自然观察，变成含有一定阶级意识因素的民约、礼仪式的社会活动，并有了氏族或部落首领之类的主试人。原始氏族社会中的"冠礼"、"洗礼"，就是当时接受青年人为氏族或部落正式成员所举行的公开而庄严的考试仪式。此种考试已经染上了阶级的色彩。

当原始社会被分工及由此导致的社会分裂为阶级而炸毁，人类进入阶级社会后，社会生产力的进一步发展，"使农业和工业之间的更大规模的分工成为可能，从而为古代文化的繁荣，即为希腊文化创造了条件"，①在东西方一些文明古国中，古文字有了新的发展，学校教育的内容得以充实，制度日趋完善，有了较为完备的国家机器。与之相适应，考试从此打上了阶级的烙印，成为上层建筑的一个重要组成部分。它由原始社会那种不成文的民约、礼仪式的考试发展为成文的法规式的考试，并分化为两种性质、内容、形式各异的考试，分别服务于不同的阶级。一种是为统治阶级服务的成文法规考试。其对象为统治阶级子女，上等自由民子女，内容是统治阶级规定的经典知识、统治之术及军事技能等，目的在于培养、选拔为统治阶级服务的御用人才。另一种是民间师徒授业考试。对象为手工业者、农民、奴隶及下层自由民，内容是生产技艺，目的在于获得社会承认、争取谋生的权利。两种考试的方式、方法也各不相同：官方考试由政府官员主持，按照统治阶级的意志，根据既定的规程，在特定的场所以面试（偶有笔试）进行。民间考试的主持者是行会首领或师傅，一般是在临时选定的场所进行，测试方法多为实际操作或演示，兼以问答。此后，由

①《马克思恩格斯选集》第3卷，人民出版社1972年版，第220页。

于社会生产力水平和文化科学水平的不断提高，社会对劳动者素质与智能的要求始终处在发展变化中，考试的类别，形式、内容、程序、方法、手段等，也一直在变化发展，但论其性质，都无不具有鲜明的阶级性。在阶级社会里，考试的领导权为掌握生产资料的统治阶级所控制，不同社会或同一社会不同历史阶段的统治者，总是根据本阶级的意志和利益制定考试制度，颁布考试法规，建立考试的组织机构，规定考试的对象和内容，确立育才、选才、用才的标准，借以网罗所需人才，达到维护和巩固其统治的目的。因此，在奴隶社会、封建社会和资本主义社会中，考试是剥削阶级统治人民的工具，其剥削阶级的特性客观存在。当无产阶级夺取政权、建立了生产资料公有制后，考试产生了质变，它由统治劳动人民的工具变成了维护、发展社会主义生产力和生产关系的重要工具，成为培养、选拔、任用社会主义建设人才的基本手段，成为调动广大劳动者积极性、尽其才能于祖国事业的有效措施。所以，社会主义社会的考试，具有工人阶级的特性。

再者，考试既是一个发展过程，也是一个认识过程。从根本上说，考试的对象是人，同时又是受过专门训练的人——主试者，根据一定社会的需要，选择一定的内容、场所和方式，对应试者进行测度或甄别，它必然会受到人的发展规律的制约。而"在阶级社会中，每个社会成员都作为一定阶级的人而存在，都在一定的阶级斗争的条件下生活。"①其身心发展除生物因素的作用外，社会因素的影响极大。也就是说，人的形成和发展是生物因素和社会因素的辩证统一，在不同社会和不同阶级中生活的人，他们的思想行为，有着程度不同的阶级色彩，并必然在考试中反映出来。比如为什么考？考什么？都是受一定阶级意识支配的。

①刘少奇：《论共产党员的修养》，人民出版社1980年版，第2页。

考试学实质上是研究考试过程规律的科学。考试过程又是沿着社会和人这两种发展过程交互作用而不断向前的演进运动，所以考试具有社会和自然的双重性质，其中社会性质便决定了考试学的阶级性。考试学如同其它社会科学一样，它不能超脱一定阶级的立场、观点和方法。不同时代的人们，必将运用不同的立场、观点和方法去研究考试，其学说必然要维护不同阶级的利益，反映不同阶级的要求，从属于不同阶级的政治，服务于不同阶级的考试实践。我国的考试学是社会主义的考试学，是站在工人阶级的立场，以马列主义、毛泽东思想为指导，从中国的客观需要出发，总结我国历史的和现实的考试经验，探索和揭示考试的规律，使之服务于社会主义的考试实践，它同样具有鲜明的阶级性。

必须指出，考试还有不属于上层建筑范畴的因素。比如考试的分类、设科，考试内容中的自然科学知识和生产技术知识，考试的程序、形式与方法等，并不受制于政治与经济。因为人的身心发展和认识过程中的生物因素是自然属性，它不随政治经济的变化而变化；考试分类、设科，是社会生产力发展的客观要求，不为人的主观意志所决定；考试的形式、程序和方法，是人类长期考试实践经验的积累和认识的结晶，可供一切阶级使用，为各个阶级服务，因此，不具有阶级性。

此外，考试有相对的独立性。它的发生、发展有其内在的联系和固有的规律。考试由自然观察发展为面试再到笔试，由非文字考试到文字考试，由个别考试到团体考试，由主观性考试到客观性考试，直至今天的面试与笔试，文字考试与非文字考试、主观性考试与客观性考试的有机结合，都是考试自身发展演变的结果，而不是某种阶级意志的产物。

还需提及的是，人类社会特有的考试现象，它不单受制于社会发展和人的发展规律，还受人口结构、传统习惯、民族风俗、自然环境等多种因素的影响，这些因素中也含有无阶级性的成分。

基于上述原因，作为反映考试现象并揭示其本质的考试学，又因考试自然属性的一面而兼有自然科学的性质。事实上，社会科学与自然科学之间也是不可截然分离的，人类科学是一个巨大的系统网络，如德国著名物理学家麦克斯·普朗克指出的那样："科学是内在的统一体，它被分解为单独的部门不是由于事物的本质，而是由于人类认识能力的局限性。实际上存在着从物理到化学，通过生物学、人类学到社会科学的连续链条。"①考试学是在现代教育、心理、管理、生理、行政、人才、统计、电脑等科学的不断发展和日益渗透的基础上，产生的一门新兴学科，它含有社会科学和自然科学两种性质。

同时，考试学还是一门应用性很强的学科，具有强烈的实践性。人类的考试实践，既是考试学理论的源泉，又是检验考试学理论的客观标准；既是建立考试学理论体系的土壤，又是考试学理论的必然归宿。考试科学理论体系的形成，必以符合人与社会发展客观规律的考试实践经验的高度概括和科学总结为基础，并在长期服务于考试实践的过程中不断丰富、发展和完善。没有人类漫长的考试实践，考试学便是无源之水；脱离考试实践，考试学也就失去了存在的前提和价值。来源于实践，服务于实践，受检于实践，是考试学的基本特性。

三、考试学的原则与方法

在科学研究中，方法犹如过河之舟。一门学科倘若没有符合其特点的研究原则和方法，即使有固定的对象、丰富的素材、充实的内容，也不可能达到理想的彼岸，成为真正的科学。固然，任何一门科学的研究，都不免遇到现象与本质、原因与结果等共性问题，在分析和释义过程中，所采用的方法不无相同之处。但是，每一门科学只能是研究客观世界发展过程的某一阶段或某一

①《新技术革命与管理》，湖北科技出版社1984年版，第14页·

运动形式，各有其特定的对象，因而也就需用不同的研究方法去认识和揭示现象与本质间的差异，使那些隐含在现象内部的固有规律实现外化。

考试学正值创建时期，作为一门独立的学科目前尚未完全成熟，还没有公认的科学研究原则和方法可资遵循。根据科学研究一些带有普遍性的原则和方法，结合考试学的特定对象和需要探讨的特殊问题，可采用下述原则和方法进行考试学的研究。

（一）考试学研究的原则

思想性与科学性统一的原则，是研究考试学的首要原则。贯彻这一原则，关键在于坚持正确的思想指导和实事求是的科学态度。无论哪门学科，也不管它以什么为研究对象，都不能没有某种哲学思想的指导。所不同的是，究竟以什么哲学思想作为自己的理论基础，要取决于研究者的立场和态度。考试学需要在科学总结和概括人类考试实践经验的基础上，建立起符合事物本身发展逻辑的理论体系，这就不能不借助思维工具，如达尔文所说，要认识本质，只能通过思维，必须历经科学抽象的两个进程，即“在第一条道路上，完整的表象蒸发为抽象的规定，在第二条道路上，抽象的规定在思维行程中导致具体的再现。”①在考试学的研究中，要把这两个进程有机地结合起来，实现从第一个具体（感性的具体）到第二个具体（理性的具体），使考试学成为一门正确反映考试活动规律的科学，就需坚持以辩证唯物主义和历史唯物主义的哲学思想为指导，以求实的科学态度，去考察、分析人类社会的考试实践活动，把考试学的研究立基于马克思主义方法论。

考试是随人类社会的产生而产生、发展而发展的，在其漫长的演变过程中，它不断地由量变引起质变，又因质变导致量变。

①《马克思恩格斯选集》第2卷，人民出版社1972年版，第103页。

考试不可能始终保持旧质而无新质代起，同时，代起的新质在它新的量变过程中，也不是毫无旧质中科学成分的保留。况且考试涉及范围宽广，纵有前后历史的承袭，横有政治经济、文化科学等方面的联系，不同历史时期的考试理论、制度、内容和方法错综复杂，常常是"精华"与"糟粕"交织在一起。考试学要从考试发展演变的悠久历史中，从考试与内外极为繁复的关系中，探明其规律性，实在不易。"为了真正获得正确处理这个问题的本领而不被一大堆细节或各种争执意见所迷惑，为了用科学眼光观察这个问题，最可靠、最必需、最重要的就是不要忘记基本的历史联系，考察每个问题都要看某种现象在历史上怎样发生，在发展中经过了哪些主要阶段，并根据它的这种发展去考察这一事物现在是怎样的。"①只有运用辩证唯物论和历史唯物论的观点，认真分析这些基本历史联系的事实，才能从本质上弄清考试变化发展的成因。同样，我们也只有运用马克思主义关于经济基础与上层建筑的关系及一切事物都是彼此联系的观点，去研究考试与政治经济、文化科学、人的发展之间的辩证关系，才能真正认识考试在人类物质文明和精神文明建设中所起的作用。

有了正确的思想导航，还必须有科学的研究态度，即在弄清客观事实、占有可靠研究素材的前提下，本着实事求是的原则，坚持具体问题具体分析，不可随意肯定或轻易否定。列宁曾指出，正确观察、分析问题的方法，应是从事物产生的历史开始的，"要正确地认识它，要有把握地切实地解决它，就必须从历史上把它的全部发展过程加以考察"，②然而，我们"判断历史的功绩，不是根据历史活动家没有提供现代所要求的东西，而是根据他们比他们的前辈提供了新的东西"。③应把他们的实践与

①《列宁选集》第4卷，人民出版社1975年版，第43页。

②《列宁选集》第4卷，人民出版社1975年版，第44页

③《列宁全集》第2卷，人民出版社1959年版，第150页。

理论放到相应的时代中去考察正误、分辨真伪。但是，又不能因为人们创造自己的文化，"不是在他们自己选定下来的条件下创造，而是在直接碰到的、既定的、从过去继承下来的条件下创造"①的历史局限性，而脱离客观实际地夸大其科学价值或历史作用。否则，就不是实事求是的科学研究态度，也不可能真正做到思想性与科学性的统一。

其次，研究考试学务必坚持理论联系实际的原则。考试学的研究与考试的实践密不可分，因为考试学所要研究的一切课题都寓于考试实践之中，而考试学研究的结果又将付诸于考试的实践。从这个意义上说，没有考试实践，就没有考试的理论；没有科学的考试理论指导，也就不可能有科学的考试实践。

考试学的研究之所以必须与实际相结合，还在于考试学自身成熟、发展和完善的客观需要。一方面，考试学的理论只有坚持从实践中来，即在研究过程中参与考试的设计和实施，运用考试学理论分析、解决考试中的实际问题，开展不同形式和规模的测试实验等，才可做到从实际出发，提高理论的实用价值。另一方面，让考试学的理论回到实践中去，检验理论的价值，看其是科学的、富有实际意义的，还是幻想的、与实际相冲突的，并通过实践的验证发现新问题、解决新矛盾、总结新经验，开展新的实验和研讨，进而实现理论上的推陈出新。这无疑是一条深化考试学研究、发展和丰富考试学理论的必由之路。

第三，考试学的研究，还应坚持"古为今用"、"洋为中用"，批判继承与择善吸收的原则。我们必须承认，"无产阶级文化应当是人类在资本主义社会、地主社会和官僚社会压迫下创造出来的全部知识合乎规律的发展。"②考试是人类文化的一个组

① 《马克思恩格斯选集》第1卷，人民出版社1972年版，第603页。

② 《列宁选集》第4卷，人民出版社1975年版，第348页。

成部分，在其理论、类型、内容、程序、方法等方面，都有一定的历史继承性，今天的考试是前天与昨天考试的继续和发展。作为“马克思主义的历史主义者，我们不应当割断历史”①，应该看到，历史上许多优秀的考试思想和成功的考试经验，它是前人智慧的结晶，也是后辈珍贵的精神财富，不同时代的人们，都可吸取滋养于其中。在科学技术飞速发展、人类考试急剧变革的今天，为使考试学成为反映考试客观规律的科学，能真正促进考试的改革和正确指导考试的进一步实践，就应象马克思那样对待人类文化的遗产，“凡是人类社会所创造的一切，他都用批判的态度加以审查，任何一点也没有忽略过去。凡是人类思想所建树的一切，他都重新探讨过，批判过，在工人运动中检验过”②。有选择地吸取历史上进步的、带有规律性的精华，扬弃反动的、违背科学的糟粕，借以开阔研究思维、丰富研究内容、得出科学结论，使考试学的研究达到承上启下、继往开来之目的。我们还应看到，考试作为一门科学，它并不受限于某个国家和民族。马克思主义历来认为，科学是社会发展的共同产品、人类共有的精神财富，“它包括任何时期和任何地区的精神产品，而不局限在某个时期、某些地区或民族的范围内。”③考试学研究对象的范围纵贯古今、横涉中外，它要揭示的是人类社会考试活动的规律。因此，在研究中必须放眼看世界，“应当尽量吸收进步的外国文化，以为发展中国新文化的借镜”④。对于国外历史上特别是近

①《毛泽东选集》（合订本）人民出版社1971年版，第499页。

②《列宁选集》第4卷，人民出版社1975年版，第347页。

③［南斯拉夫］费拉基来尔·鲍皮克：《关于教育与教养的理论研究方法的探讨》，载《华东师范大学学报》1986年第1期，第4页。

④《毛泽东选集》（合订本），人民出版社1971年版，第984页。

十几年来考试实践中不失科学价值的经验，一些被实践证明符合规律的理论，以及某些先进的考试手段和方法，我们本着"洋为中用"的原则择善吸收，定将有益于考试学的研究。

第四，实现考试高度数量化，是当今国内外考试界竭力奋斗的目标，也是考试科学研究中的重大理论课题。因之，定性分析与定量分析相结合的原则，为考试学的研究必不可少。

人类社会是不断进化的，它从原始的渔猎社会发展为农业社会，又历经工业社会而步入当今的信息社会。与此相一致，科学技术的发展和学科间相互渗透的速度日益加快，人们认识客观事物的途径、分析各种现象的方法和手段，也在不断地改进和更新。过去社会科学领域那种单一的定性分析法，现已为定性与定量相结合的综合分析法所取代，社会科学同自然科学在思想与方法方面，相互渗透、汇流的现象日趋明显。考试学是一门带有综合性的学科。数理理论是它的科学基础之一，无论是考试模式的建立、程序的编制、实施过程的控制，还是考试结果的分析和处理，除部分使用定性分析外，大都要借助数理统计方法，进行量的分析、描述、计算和推导。考试设计、实施、直至结果分析处理的全过程，实际上是以量化为主导的控制过程。作为反映这一过程内在规律的考试学，倘若仅有定性分析的手段，显然研究目的难以期及。考试学的成熟，数学和电脑这一"科学的得力助手和工具"非有不可，正如保尔·拉法格在《回忆马克思》一书中所指出，一种科学要想真正达到完善的地步，在很大程度上取决于它能否成功地运用数学。

（二）考试学研究的方法

一门学科研究方法的选择与确定取决于两点：它所遵循的理论原则和所要研究的具体对象。由于辩证唯物主义和历史唯物主义是考试学的方法论基础及其研究所遵循的基本原则，人类社会的特有现象——考试，是考试学研究的具体对象；考试学的理论

源于实践又回归于实践，如若割断彼此间的联系，无疑会丧失考试学的生命力。因而，在马克思主义关于客观性、阶级性、历史与发展联系、理论与实践统一的认识原理指导下，调查研究、统计分析、比较、测试实验等，应是考试学研究的基本方法。

1. 调查研究法。

考试学归根结底是从客观考试事实中引出的规律性结论。它的研究必须充分占有材料，全面了解和掌握客观事实。为此，就需通过调查研究的方法来解决。调查研究的途径和方法多种多样，大致可分为两大类：直接调查与间接调查。

直接调查的基本特点是深入实际，以人为主要调查对象，获取真实可靠的第一手感性资料。在调查过程中，看、听、问、作等方式兼有，一般采取研究者直接参与考试的设计和实施、考试结果的分析与处理；亲临施测现场观察，与主试者、被试者接谈；对经过测度、甄别而入选、录用、晋升的对象，进行长期或定期、全面或重点的观察；走访育人用人单位及被试家庭；出国参观、考察等。

间接调查则以文字材料为主要调查对象，内容包括历史的、现实的、系统的、零碎的。所获取的有第一手资料，也有第二手资料。调查方式不如直接调查复杂，除少量考试录像、录音需看、听结合外，基本上都是与文字资料打交道。间接调查通常以查阅资料和问卷两种方法进行。问卷法亦称书面调查法，即以书面形式提出所要调查的内容，让被调查者用文字、数表、图型、符号等形式作答，以了解考试的利弊、社会效用、心理反映、问题及改进意见等。资料查阅是考试学研究的主要方法之一。它通过查阅历史文献、文书档案、地方志、统计资料等，得到诸如不同社会及历史时期考试法规、考试组织机构及其职能、考试内容与形式、考试方法与手段、主试与被试的条件、考试的投资与经济效益等方面的第一手资料。此外，还可通过查阅有关史籍、教

科书、报刊等，获取第二手资料，以了解各个时期不同国家、民族、区域及社会不同方面有关考试的记述和见解。

在考试学的研究中，无论采用何种调查手段，也不管所取得的是第一手资料还是第二手资料，目的都是为得出考试现象的规律性结论而搜集素材。在调查研究时，切忌材料笼统、事实含糊。所集材料必须客观、典型、面广、量足，要"从事实的全部总和、从事实的联系去掌握事实"①，而不是片断地、随意性地猎取。

2．统计分析法。

统计分析是调查研究的延伸。如果说通过调查了解客观现象是科学研究入门的向导，那么，科学的统计分析便是打开理论大门的钥匙。换言之，仅有调查没有统计分析，人们对客观事物的认识就不能由感性上升为理性，真正透过现象把握实质。

科学的统计分析应是定性与定量分析相结合，需有分析、综合、归纳、演绎、类比、统计、推导等多种方法的参与。科学分析的任务，在于对调查所得各种感性材料去粗取精，去伪存真，由此及彼，由表及里，通过整理、概括和总结，使之条理化、系统化等加工改造，将感性的经验提炼成理性的认识，进而上升为科学理论。整理分析调查材料尤其要注重数理统计，这是考试学研究考试可控原理的重要手段，也是检验考试质量的基本方法。在素材的加工整理过程中，只有正确使用统计方法，为研究提供各类可靠数据，才能具备进行因素分析和质量分析的条件，才能探明考试"群体"诸种关系的实质，达到揭示规律的目的。

3．比较法。

有比较才能鉴别，这是对人们认识客观事物方法的规律性总结。比较作为一种判定事物异同关系的方法，它是分析、抽象、概括、归纳、演绎等多种思维形式的综合运用。通过不同层次的比较，可使人们对同一事物达到不同程度的认识。一般说来，采

①《列宁全集》第23卷，人民出版社1958年版，第279页。

用类比的方法可判别事物的同异之点，并将事物按其属性归类。若采取对比的方法（包括相反的同一属性比较，近似的同一属性比较，同一属性因素与数量的比较），亦即将不同事物内部矛盾的一切方面进行对比分析和综合论证，便可区分事物质的异同。

比较法有多种类型，如分类比较、综合比较、垂直比较、水平比较、正反比较、近似比较、程度比较、因素比较等。在实际研究中，人们往往因研究对象的需要有选择性地应用。根据考试学所要研究的内容，采用垂直比较、水平比较和因素比较的方法较为适宜。

垂直比较，或称纵向比较，即把人类历史上不同社会的人才观、教育观、教学观、考试观、考试制度、考试组织系统、考试实践等，进行分门别类的比较，以分辨不同社会考试的特点、模式结构的异同，弄清各社会间考试的前后关联，进而找出考试发展演变的内在规律。

水平比较，又称横向比较。它是将同一社会或同一历史时期不同国家、地区的考试，进行全面对比分析，内容涉及考试法规、考试类别、考试的方式方法、考试手段、考试程序、考试内容、考试理论、考试机构与管理、考试的实际效果等。意在总结不同国家和地区考试的共同特点与变化趋向，然后，再通过综合比较和系统分析，寻求考试活动的基本规律所在。

因素比较，它是将同一属性的问题作定性与定量分析，通过利弊优劣的比较研究，找出考试中不同问题的各种影响因素何在，判明考试与社会生产力、科学文化技术发展、人才、教育、人的身心发展，以及考试内部各环节之间的相互关系，帮助我们透过这些纷繁的联系现象，认清事物的本质，作出科学的结论。

比较是科学研究的主要方法之一。马克思、恩格斯在《德意志意识形态》一书中有过高度评价，他们认为比较研究法具有普遍意义。毛泽东同志也曾有"没有比较，就没有鉴别"的名言。但使

用比较研究法必须有严谨的科学态度。比较的过程，应是在马克思主义方法论指导下，按照客观标准实事求是地考察和分析问题的过程，是探求规律、寻找正确答案的过程。如果东拼西凑、随意作比、草率定论，或仅有现象的比较而无本质的分析，流于一般性事实叙述，都会失去比较的意义。

4．测试实验法。

在考试学的研究中，通过调查研究、统计分析和比较等方法，我们可以找出对于科学认识考试确有价值的客观事实，明了某些考试现象的内在机制，揭示出考试活动的基本规律。或者说，获得了考试学这门学科所要论证和表述的东西。但这并不说明所得出的结论都是绝对科学的。采取这些方法概括、总结出来的理论正确与否，必须经过社会生活检验后，方可定论。因为"社会生活在本质上是实践的。凡是把理论导致神秘主义方面去的神秘东西，都能在人的实践中以及对这个实践的理解中得到合理的解决"。①考试的实践，是对人的素质与智能进行测度或甄别的社会活动，通过实际测试是验证考试理论最基本的途径，也是最有效的检验方法。

测试实验法，是根据研究者所提出的理论，有目的、有针对性地设计出类型有别、内容各异、对象不同的考试方案，严格按照考试实施程序进行实际测试。然后，对测试结果作全面细致的统计分析，从而得出所用理论正确性与可行性如何的结论。形式多样、规模不等的测试实验，不单是为考试理论的科学与否寻求客观依据，以便推广和应用，同时也是为考试科学理论的发展和完善充实新的内容。没有广泛而持久的测试实验，就不能形成富有中国特色的考试科学理论体系。

综上所述，考试学的研究原则和方法，虽应遵循科学研究一些共同的认识论原则和方法论，但因其研究的内容与任务的特殊

①《马克思恩格斯全集》第3卷，人民出版社1972版，第5页。

性，更多的是用本门学科的原则与方法，并根据研究问题的不同性质和特点，将上述各种原则、方法有机结合，恰当地交替使用。上文所列原则和方法，其独立性是相对的，彼此逻辑联系紧密，是一个互为关联的体系，不能彼此割裂。

考试学的研究方法，可归纳为经验方法与理论方法两大类。但是，这两类研究方法究竟应包括哪些具体的方法才真正适合考试科学研究的需要，仍是有待深入探讨的问题。

第二节 考试学的科学依据

如前所述，考试学的产生并非偶然，它是现代社会生产与生活的客观需求、现代科学进步的结果、考试自身长期演变发展的必然。考试学这门新学科的创建，诚然也不是无中生有、凭空而论，其科学依据是客观存在的。

那么究竟什么是考试这门科学的依据呢？要弄清这一问题，首先应明确"科学"概念的含义及其与客观事物的关系。所谓科学，概要地说，即是人们认识和改造世界实践活动的结晶，或曰人们正确反映世间诸种事物发生、发展之客观规律的知识体系。它同客观事物之间，是认识与被认识、反映与被反映、揭示与被揭示的关系。在唯物论的认识观看来，"人的认识物质，就是认识物质的运动形式"，①"观念的东西不外是移入人的头脑并在人的头脑中改造过的物质的东西而已"。②倘若不存在被反映之物，即没有客观存在的运动的物质，以及每种物质运动形式内在固有的特殊矛盾，也就无所谓各门科学的存在。因为，科学的任务不外乎揭示客观事物的必然联系和各种客观事物发展过程的内在规律性，所

①《毛泽东选集》第1卷，人民出版社1971年版，第283页。

②《资本论》第1，人民出版社1986年版，第24页。

以，科学研究首先必有客观存在的对象，然后才能按其研究对象的本来面貌去考察和认识。否则，科学就无所依存，也不可能成其为科学。正是根据唯物论的认识观，毛泽东同志明确指出："对于某一现象的领域所特有的某一种矛盾的研究，就构成某一门科学的对象"①，而每一客观事物运动形式所包含的特殊矛盾，便规定了一事物区别于它事物的特殊本质，进而也就成为研究该事物固有的必然联系及内在规律之科学的依据。可见，一门学科有无客观存在着的特定研究对象，当是判定一门学科是否有科学依据的关键。

考试学的研究对象，是测度、甄别人的德、学、才、识、体个别差异的社会活动——考试活动，它所要揭示和反映的是考试现象固有的必然联系和运动规律。既然考试现象不为人的意志所决定，而是从根本上受客观规律支配的，那么，反映考试现象客观规律之考试学的依据，也必为其研究对象所具有的普遍的和特殊的矛盾性所决定。作为社会现象的考试，是独立于人的主观意识之外的客观存在，它通过感觉、理性的形式反映到人们的主观意识中，因而，它具有一切物质共同的最根本的特点或普遍的本质。"而承认外间对象的客观存在性，按照研究对象的本来面貌去认识它们，又正是唯物论哲学的基本要求"②，也是一切科学研究理应遵循的普遍原则。任何一门学科对其研究对象规律性的揭示，都必须用与现实严格一致的抽象概念来表达这些规律。考试学的研究对象所具有的这种普遍的矛盾性，便规定了考试学应有与其它学科共同的普遍性依据——马克思主义唯物论。考试学只有在唯物论哲学观的指导下，运用马克思主义的科学方法论，按照考试现象的本来面貌，去考察、研究考试的发生、发展及其

①《毛泽东选集》(合订本)，人民出版社1971年版，第284页。

②李达：《唯物辨证法大纲》，人民出版社1978年版，第24页。

各方面的联系，才有可能建立并形成正确反映考试规律的理论体系。但“尤其重要的，成为我们认识事物的基础的东西，则是必须注意它的特殊点”①。考试学研究的对象，虽具有某些其它事物共同的特点，但更主要的是具有区别于它事物的特殊矛盾性。这不仅规定了考试现象的特殊本质，而且构成了建立考试学的特殊依据，即考试是人类社会的客观需求，人的德、学、才、识、体的个别差异客观存在，人的德、学、才、识、体的个别差异可测，考试流程有序、可控。现分述如下：

一、社会对考试的需求客观存在

一门科学的产生和发展，原因是多方面的，既与作为方法论基础的唯物论哲学的发展密切相关，又同人类科学技术的进步紧密关联。然而，最根本的是决定于物质生产和社会发展的需求。从这个角度讲，社会对考试的客观需求，乃是考试学的社会根基。

人类社会进化的历史表明，人才是振兴社会之本。无论在什么社会，人才都是最珍贵的财富，都是推动社会发展的基本动力（因为人是生产力的本质要素）。管子曾云：“一年之计，莫如树谷；十年之计，莫如树木；终身之计，莫如树人。”②一个国家或一个民族，如若没有足够数量的、具备与其社会物质生产要求相应素质和智能的人才，要想实现改变现状、发展文化、推动社会前进的目的，实不可期及。

众所周知，一个社会若无教育，新生一代将成才无道；有教育而缺少测度、甄别人的德、学、才、识、体个别差异的考试，不仅人才质量优劣不可辨，而且难以庶全盘人才作统筹有效之用；若有育才、选才之良方，而任使无科学之法，同样不可使已

①《毛泽东选集》第1卷，人民出版社1971年版，第283页。

② 管仲:《管子·权修》。

有的人力发挥应有的社会效用。任何社会要不乏所需人才，并收到才尽其用之效，就必须有教、考、用三者的密切配合。有所区别的是，由于人类物质生活的生产方式始终处在变化状态，而"物质生活的生产方式制约着整个社会生活、政治生活和精神生活的过程。"①在物质资料生产方式的各因素中，又以生产力最为积极活跃，一切社会关系的变化，一切社会活动的内容、速度和规模，都取决于生产力的发展水平。考试和教育均属人类有意识、有目的的社会活动，它的发展是与生产力的发展水平相一致的，其内容、规模及其结构，必受生产力发展水平的制约，所以，在生产力水平、经济、技术和产业结构不同的社会或历史阶段，社会对人才质量的需求是不尽相同的，进而对考试和教育的要求及其标准也就不一样。原始社会时期，限于生产力水平的低下和文化科学的贫乏，教、考、用三者不必要也没有可能独立，一切都在生产劳动和日常生活的过程中进行。随着生产力的发展和体力劳动与脑力劳动分工的出现，社会产生了专门掌管国家事务、从事文化科学研究活动的机构。于是，需有一批具备相应素质与智能的新型人才行使其职能。社会由此对考试有了新的需求，即根据社会和统治阶级的要求甄选欲培养的对象，经过专门教育机构——学校授予他们即将从事职业的知识与技能，再通过测度或甄别，选出统治阶级所需的各种官吏，以维护和巩固其统治。同时，也是社会持续和发展的客观需求。当纺织机、蒸汽机敲响了现代生产的钟声，人类由农业社会跨进工业社会后，劳动过程的社会结合不断加强，分工越来越细，新的生产力要求劳动者不断提高文化素质，具有较强的读、算技能，懂得并能正确使用机器。社会对数以百万计的劳动者素质与智能的新要求，便赋予教育和考试新的使命。一方面，要求学校用"批量生产"的现代培养方式取代师传徒的培养方式，以加快智力资源的开发，提

①《马克思恩格斯选集》第2卷，人民出版社1972年版，第82页。

高教育投资的效益。另一方面，要求考试调整结构，向多样化、系统化发展，充分发挥其教学功能与社会功能，即按照社会的需求和统治阶级确立的标准，挑选入学对象、检查学习效果、衡量教育质量；甄选各级各类官员、招收雇用劳动力，定期不定期地评量在业人员的素质与智能，以调整各生产部门的人才结构，实现人力资源的充分利用和生产效率的成倍提高。至此，社会对考试的需求也并未终了，因为“现代工业的技术基础是革命的”，它“从来不把某一生产过程的现存形式看成和当成最后的形式。……现代工业通过机器、化学过程和其他方法，使工人的职能和劳动过程的社会结合不断地随着生产的技术基础发生变革”。① 经过19世纪末和20世纪中期两次技术革命，现代生产终从蒸汽机起步，走完了电气化的历程，进入了以电脑技术、生物技术、新材料、新能源为主要特征的新质阶段，形成了现代化科学知识和智力劳动为基础的生产，并“不断地使社会内部的分工发生革命，不断地把大量资本和大批工人从一个生产部门投到另一生产部门”②，从而导致社会技术、产业、企业、职业结构不断地发生变革，使“农业，这个文明世界所赖以生存的最早的基础，在一个又一个国家里失去了自己的支配地位。”③第三、第四产业迅猛发展；在一些发达国家，脑力劳动者在从业人员中的比例已超过体力劳动者；各生产部门由劳动力密集型渐变为知识密集型；从前那种“以人的体力扩展和强化为基础”的技术革命特征不复存在，强化脑力和削弱体力成了当代技术革命的主要特点，智力和知识已是构成现代生产者劳动技能的基本要素。由此，社

①《马克思恩格斯全集》第23卷，人民出版社1972年版，第533、534页。

②《马克思恩格斯全集》第23卷，人民出版社1972年版，第534页。

③〔美〕阿尔温·托夫勒：《未来的震荡》，四川人民出版社1985年版，第10页。

会对劳动者的素质与智能提出了人类历史上不曾有过的要求。在蒸汽机时代，要求劳动者接受初等教育；在电气化时代，要求劳动者具有中等文化水平，而今，则要求劳动者接受中、高等专业教育和终身教育，而且仅有大脑蓄积的死知识已远不能适应社会发展的需要，必须牢固地掌握宽厚的基础理论知识，以及各种知识之间的联系，凭借一套运用知识的"参照架构"，去有效地驾驭和灵活地使用所学知识，并不断发展自学能力、研究能力和创造思维能力，才能适应不断变革的现代生产的需要。

不难看出，在技术革命浪潮的冲击下，新的生产部门的出现、老的生产部门的淘汰，劳动者从一个部门转换到另一部门，从而引起社会内部分工、企业结构、职业结构、生产者智能结构的变革，这正是现代生产发展的基本规律。教育和考试的功能，也因此而不断地得到强化。在现代社会中，教育已成为人们开发智力和"学会生存"的主要手段和发展社会经济的基础。与此同时，社会对劳动者空前未有的新要求，又迫使测度、甄别劳动者德、学、才、识、体个别差异的考试改变传统模式，冲破偏重知识、记忆能力测量的狭小圈子，达到全面、客观、准确地测度劳动者知识、技能、能力、品质、体能等的新水平，既要为亿万劳动后备军的智能发展导以正确方向，促进现代教育目标的实现，又要准确有效地选拔社会各部门的所需人才，做到学用对口、能级相称、才尽其用，实现人才培养、甄选、使用、管理科学化。与社会的需求相适应，科学的考试，已成为现代社会实现教育目标的"整速器"，调节现代生产部门人才结构的有力杠杆，开发、利用人力资源和促进现代生产发展的重要手段。

马克思在《资本论》中曾断言，现代生产过程的现有形式绝不是最后的形式。生产者的职能，劳动过程的社会结合，社会内部的分工，社会对测度、甄别劳动者德、学、才、识、体个别差异的考试的需求，都将随生产力的发展、科学技术的进步而不断

地发生变化，鞭策人们根据新的社会需求，去研究和革新考试的理论、内容、方法与手段。只要科学技术的发展，以及社会内部的分工、生产者智能结构、劳动过程的社会结合的变革不终止，社会对判别人的个别差异的需求，就必然客观存在，考试也就不会从社会生产和社会生活中被清除，人们对考试规律的认识和探索，也将无止境。所以说，社会对劳动者德、学、才、识、体个别差异测度、甄别之需求的客观存在，是考试学产生的社会根基，也是考试学赖以发展、成熟的沃土。

二、人的德、学、才、识、体的个别差异客观存在

人是自然存在物，又是社会存在物。由于每个人身心发展的物质前提——生来具有的自然资质不尽相同，身心发展的决定性条件——后天生活的社会环境与教育各自有别，所以，每个人在德、学、才、识、体方面具有差异性。这种个别差异的客观存在，以及关于个别差异的理论，便是考试学的基石。

人的素质，是个人生来具有的生理解剖上的特点，主要包括人的肌体的结构与形态，感觉器官、运动器官、神经系统、大脑结构及其机能的特点。它只是人的身心发展的物质条件，不决定人的心理发展的内容和水平。人的心理来源于社会实践。人的素质也是在社会实践中不断发育和成熟的，它随有机体年龄的变化由弱至盛再到衰。

智能是人的属性，属个性心理学范畴。人们通常用智能统称人的智力和能力。智力和能力二者相互交叉，能力是智力的表现形式，智力存在于能力之中。但智力与能力两项指标并非成正比，智力水平高的人，其能力和创造效果并不一定都高。智能又与人的知识、技能密切关联，智能的发展可促进知识、技能的掌握，知识、技能的掌握，又是智能发展的重要途径和条件，两者互为依存、相互促进、统一实现，既不能互相割裂，也不能彼此

混淆。知识是人对客观事物的认识，是人类历史、经验的总结，是推动科学技术发展、提高生产力的力量。技能是掌握并应用某种专业技术的能力。概言之，智能是人的独立分析问题和解决问题能力与智慧的总称。

人类长期的社会实践和理论研究的结果证明，无论是人的生理素质，还是知识、能力、智力、技能、技巧、品质和体能，其个别差异都客观存在。中外历史上不同时代的教育家、思想家、心理学家、生理学家和医学家，对人的体质、体能、气质、性格、智力、技能、能力个别差异的论述颇多。早在春秋战国时期，我国伟大的教育家孔子，就已认识到人的个别差异的存在，他明确提出了"性相近也，习相远也"①的个别差异理论。宋代朴素唯物主义者张载，对人的个别差异的论述，较之前人更为全面深刻，他在《张子语录》中写道："人与植物之类已是大分不齐，于其类中又极有不齐。某尝谓天下之物无两个有相似者，虽则一件物亦有阴阳左右，譬之人一身中两手为相似，然而有左右，一手之中五指而复有长短，直至于毛发之类亦无有一相似。至如同父母之兄弟，不帷其心之不相似，以致声音形状亦莫有同者，以此见直无一同者"。②他的思想对明末清初杰出的唯物主义哲学家王夫之影响殊深，其关于个别差异的理论为王夫之所深化。王夫之认为是因后天不同社会环境的影响，才造成了人的"情"、"识"、"意"等心理特征的差异，即"情识意见成乎万殊者，物之相感，有同异，有攻取，时位异而知觉殊，亦犹万物为阴阳之偶聚而不相肖也"。③王夫之在《尚书引义》中精辟地阐述了"性日生而日成"，"习与性成者，习成而性成也"的理论，认为"性"因"习"而异，人的心理差异客观存在，并伴

①《论语·阳货》。

②《张子全书》。

③王夫之：《张子正蒙注》。

随人的生命的伸延而发展变化，他说："神智乘血气以盛衰，则自少而壮，自壮而老，凡三变而易其恒"①。此外，在智力的差异方面，汉代王符在其《潜夫论·德化》中有"上智与下愚之民少，而中庸之民多"之说，指出了人在智力水平上的差异。三国时的刘劭，进一步阐明了人之才智成有早晚的差别，他说："有早智而速成者，有时晚智而晚成者，有少无智而终无成者"②。孔子、王充等人，对人的智能水平高低不一，也有明确评述。有关人的性格差异，历史上有不少真知灼见，集大成者当首推刘劭，他在孔子关于人之性格有"狂者"、"狷者"、"中行"之别的基础上，对人的性格差异作了细微深入的探讨，得出了人的性格差异有12类之分的结论。至于人的气质差异，公元5世纪希腊医师希波克拉特等人就提出了"四体"说，我国《黄帝内经》的论述更为精细。

总之，人与人之间在心理、生理方面的差异现象客观存在，有关这方面的史料可谓汗牛充栋，无庸详述，况且前人在人的个别差异方面的感性经验或理性认识，绝大部分已在现代生理学、心理学及解剖学的研究中得到印证。通过现代科学的研究，人们进一步弄清了人的生理机制和心理活动的奥妙：人的个别差异虽有先天因素——遗传素质的影响，但造成个别差异主要在于人的后天因素——社会环境各各相异所致。马克思在《哲学的贫困》中指出："搬运夫和哲学家之间的原始差别要比家犬和猎犬之间的差别小得多，他们之间的鸿沟是分工掘成的。③"恰如爱迪生所总结："天才是百分之一的灵感，百分之九十九的血汗。"鲁迅先生说得好，即使是天才，生下来的第一声啼哭，也绝不会就是

①王夫之：《读通鉴论》。

②刘劭：《人物志·七缪》。

③《马克思恩格斯选集》第1卷，人民出版社1972年版，第124页。

一首好诗。事实上，人的天赋素质的差异并不是一条不可逾越的鸿沟，某些先天因素的缺陷，是可以通过学习和实践得到不同程度的弥补的。人在德、学、才、识、体诸方面之所以参差不齐，也正是教育和社会环境起了决定性的作用。马克思主义认为，有存在才能有反映，有实践理论才有源泉。人的德、学、才、识、体个别差异的客观存在，人类测度、甄别人的个别差异的长期实践，为考试学的理论提供了被反映的客体，是考试学揭示人类社会特有现象——考试固有规律的坚实基础。

三、人的德、学、才、识、体的个别差异可测

人在德、学、才、识、体诸方面的个别差异不但客观存在，并通过语言和实践活动表现于外，既可观察、记录，又可测度、甄别。这是考试存在的前提，也是考试学的科学依据之一。

有数量必有质量，任何质量都含有相应的数量，这已是为人类实践和科学验证的结论。既然我们承认人的才能有大小之差，“智慧有高低之分，品格有优劣之别，在这高与低、优与劣之间，必定有程度的不同，所谓程度的不同，就是数量上有差异；而数量上有差异的东西，当然有测量的可能。”①对此，2000多年前的孟子就已经觉察，认为人的身心个别差异如同世上其它事物一样，是能够进行“权”、“度”的。他在《孟子·梁惠王》中阐述了自己的观点：“权，然后知轻重；度，然后知长短。物皆然，心为甚”。时过200余年，我国两汉交替时期的进步思想家桓谭，在人的生理、心理个别差异的认识方面取得了长足进步，对人的身心个别差异有了初步的量化概念，提出人的质行、才能、智略等有“大小”、深浅”、“暗照”、“薄厚”之别，他说：“若材能有大小，智略有深浅，听明有暗照，质行有薄

①孙邦正：《教育测验及统计》，台湾商务印书馆1986年第10版，第3页。

厚，亦则异度焉"①。这些论述，因历史及科学技术条件所限，不可能真正揭示人的身心个别差异的本质，更难精确地数量化。但是，我们不能因此而对古人的理论抱以讪笑，他们终归认识到了人的个别差异是可测的，并为后世指出了一条深入探究的途径。

到19世纪，德国实验心理学家冯特，终于在人的心理研究史上树起了新的里程碑，把心理科学推上了独立发展的道路，使人类对人的生理机能及心理现象的研究进入了一个崭新的时代。1859年达尔文的科学巨著《物种起源》的发表，又给了心理学研究以强烈影响。尤其是书中关于物种各成员之间的变异理论，即物种个别差异的阐述，引起了达尔文表弟高尔顿的极大兴趣。高尔顿于1882年在伦敦创设了人类学实验室，开始利用仪器作心理测量，又用他在个人心理学家克特雷关于常态误差律理论的基础上创造的数学统计方法，着手人的个别差异的研究与测量，从而使人的素质与智能差异的量化设想变成了现实，成为差异心理学的创始人。与高尔顿同时致力于个别差异研究的，还有美国的卡特尔，他们分别发表了多篇论文，相继出版了《人类才能及其发展的研究》、《心理测验和测量》等著作。他们的理论研究成果、测量手段及统计方法，被陆续应用于教育和社会其他领域，加速了教育测验和教育测量的产生和发展。此后，通过生理学、解剖学、心理学、教育测量学等近百年的研究和实验，逐步加深了对个别差异客观规律的认识，测度手段与方法日臻完善，达到了既能诊断人的生理病症，又能诊断和测度人的德、学、才、识、体个别差异的目标。

无数科学家研究结果证明，世间一切客观存在的事物，无不具有一定数量，均有被测度的可能。关于这一点，美国早期教育

①桓谭：《新论·言体篇》。

测量运动的中心人物桑代克，在其所著《精神与社会测量学导论》一书中曾明确指出："凡物之存在，必有其数量。"美国另一位测验专家麦柯尔又进一步论证："凡有数量的东西，都可以测量。"人的德、学、才、识、体的个别差异，是独立于意识之外的客观存在。这些教育的结果和精神特性，又表现在人的各种实际行为方面，"我们可以从行为方面，来测量人类精神的特性"和教育的结果，正如陈选善在《教育测验》一书中所指出："人类任何特性，凡是存在的，都可以应用已有的方法去得到较客观、较准确的度量"，只不过是度量的方法有别于物质测量而已，它采用的是间接测量法。

考试是测度、甄别人的个别差异的主要方法，经过人类长久的实践检验，证明用考试方法测度、甄别人的知识、品质、技能、智力、能力、体能等，可能且可行。近几十年来，随着人类对人的生理机能、心理反映形式及其规律性认识的不断深化，考试形式、程序、方法的进一步改革和技术手段的更新，考试所测到的质和量、主试者欲测的质和量与被试受测方面客观具有的质和量，三元合一的程度愈来愈高。科学的考试在测度、甄别人的个别差异方面的实际效果，不仅证明人的德、学、才、识、体诸方面个别差异可测的理论是科学的，而且为自身的发展开辟了广阔前景，也为考试学的产生和发展创设了先决条件。

四、考试流程有序、可控

一门学科的产生与发展，有赖于人类生产和生活客观需求这一巨大原动力的推动；一门学科的兴衰，决定于它能够满足社会需求的程度；而要使一门学科从经验的总结上升为科学的理论体系，则取决于研究的方法和手段。

在科学演进的历史上，科学的发展与科学方法论、科学手段

的发展，总是并驾齐驱的。马克思主义哲学这一具有普遍意义的方法论的产生，加速了近现代科学的发展，指导人们冲破了一个个科学的禁区，理论研究中的新发现、新成果层出不穷。而科学本身的发展又促成了科学方法论及手段的变革，在本世纪40年代便产生了"三论"（系统论、控制论、信息论）方法论。随着"三论"的成熟和发展，"人类的知识结构形成了一个'金字塔'。马克思主义哲学作为普遍抽象的科学，是这个金字塔的塔顶，中间是'三论'等横向学科，最下面是研究某一物质运动基本形式或特定范围客体的部门学科，使整个科学形成一个多层次、多结构的整体网络。与此同时，科学方法论也相应地划分为三个层次，即具有普遍意义的哲学方法论，'三论'方法论和部门学科的具体方法论。"①可见，被人们称为现代科学三大支柱的"三论"方法论，又为现代科学的发展拓宽了道路，既促进了科学技术整体化的形成，也为新兴学科的产生和发展提供了条件。考试学的基本理论体系，正是在马克思主义哲学的指导下，运用"三论"和本门学科的方法论建立起来的。

考试是一项系统工程，也是一个有规律的运动过程。它从设计、实施、直至考试结果的分析处理，犹如一条多环节的连续链条，各环节相互独立又密切关联，互为依存又相互制约，从而形成一个有序的流程。它有一套系统的科学程序和防止与测试目的无关因素干扰的措施，整个流程均在控制条件下按既定的目标运行，故考试具有目的性、有序性、可控性的特征。这些独具的特性，便规定了考试的科学基础、方法及手段，它必须综合运用现代教育、心理、生理、管理、人才、社会、行为等科学理论研究的新成果和"三论"方法论，借助现代统计方法和数学这一量化工具，以电脑技术为科学手段，也就是说，考试是通过科学理论、

①《新技术革命与管理》，湖北科技出版社1984年版，第14页。

统计方法、电脑技术三者的有机结合所构成的特殊条件，以及对这些特殊条件的科学利用，来实现考试的有序运行和对整个流程的控制的。没有确切的信息反馈、科学的统计方法和先进的技术手段，就谈不上对考试流程进行富有实效的控制。

考试的有序和可控，既是考试本身的特性，也是考试学揭示和反映考试流程运动规律的重要条件。因为科学研究的目的并非是对现象的描述，而是通过现象揭示事物的固有联系及本质规律，探求客观真理，所以，研究要有严密的程序，尤其应注重条件的控制，以便对事物的真实现象有同一反复的观察；同时，必须把事物当作一个整体或系统进行研讨，用数学模型描述、确定系统的结构和行为；还需将系统控制在相对稳定状态，尽为排除主客观因素干扰所引起的误差，使系统的功能、行为方式及其变化趋向，能在基本相同的条件下真实外化和复现，并用数理统计方法和电脑技术，对所获信息进行定量与定性相结合的分析，从事物的个性中寻求共性，又从事物的共性中发现和认识个性，即为事物的特殊矛盾性所规定的特殊本质，揭示事物的客观规律。由此看来，在科学研究中，透过事物现象找出其具有的有序性和稳定结构，当视之为正确揭示事物本质规律的基本前提。考试流程本身的有序和可控，自然为考试学认识考试系统与内部子系统及外部环境的关系，找出影响系统变化发展的因素，分辨系统内部诸环节的功用及彼此间的协同与关联，弄清整个考试流程运行规律与可控法则等，提供了客观依据，构成了考试学能够使考试现象隐含的内在本质因素实现外化的基本条件。

第三节 考试学与临近学科的关系

每一门科学都有它特定的研究对象，而每门科学的对象又决

定于它所研究事物具有的特殊矛盾性。所以，“科学研究的区分，就是根据科学对象所具有的特殊的矛盾性，……如果不研究矛盾的特殊性，就无从确定一事物不同于他事物的特殊的本质，就无从发现事物运动发展的特殊的原因，或特殊的根据，也就无从辨别事物，无从区分科学研究的领域。”①考试学之所以构成社会科学中一门独立的新兴学科，就在于它具有特定的研究对象、内容和任务，以及为此所决定的特殊研究方法和手段，其知识体系有相对的独立性。

一、学科知识体系的独立性

一门学科的独立并不意味着孤立。必须看到，世界是一个统一的物质整体，无论是自然现象还是社会现象，它们都是在相互联系、彼此制约的矛盾运动中不断向前发展的。这是自然法则。即使是同一事物，在其矛盾的统一体中，也包含着许多各具特殊性的矛盾。每一矛盾的不同方面各有其特点，从不同角度研究同一事物不同方面矛盾的特殊性，也可构成研究对象不同的学科。因此，各门学科知识体系的独立性是相对的，而学科间在知识体系上存在某些内在联系则是绝对的。同时，一门学科的研究方法和手段，虽因研究的内容和任务不同而有异，必须严格遵守马克思主义关于“用不同的方法去解决不同的矛盾”②的原则，但又为研究对象具有的矛盾普遍性所决定，需要采用带有普遍意义的方法论和原则，这就要求一门学科的研究，务必注重学科对象所具有的特殊的矛盾性，但又不可忽略与其相关或临近学科的联系。否则，就不能真正揭示区别于其他学科的特殊本质。

①《毛泽东选集》（合订本），人民出版社1971年版，第284页。

②《毛泽东选集》（合订本），人民出版社1971年版，第286页。

前文已述，考试学既是一门独立学科，又是一门综合性学科。从它的对象、内容和方法看，涉及范围极广：考试学研究的对象——考试，是人类社会普遍存在的现象，它渗透于社会各个领域的不同阶层；考试的对象是人，而"人的本质并不是单个人所固有的抽象物。在其现实上，它是一切社会关系的总和。"①牵涉政治、经济、文化、科技各方面；测度、甄别人的德、学、才、识、体个别差异是考试的本质职能，而构成人的德、学、才、识、体的要素十分复杂，其表现形式千变万化；考试学虽以社会科学的成分为主体，同时又兼有自然科学及应用技术科学的成分，其中许多问题的研究，需多种方法灵活地交替使用，并要有不同技术手段的密切配合。

由于上因，考试学便与社会科学、自然科学、应用技术科学三大领域中的许多分支学科相关。诸如，与教育学、教育测量学、教育评价学、心理学、生理学、伦理学、管理学、人才学、社会学、行政学、行为学、数学、统计学、史学、电脑科学等，在知识体系上均有程度不同的联系，还与横向科学方面的系统论、控制论、信息论有内在的关联。这种学科间的联系，正随现代科学相互独立又相互渗透、不断分化又高度综合趋势的发展而日益加强。考试学的研究合理利用这些联系，将相关学科中确有价值又必须借鉴的成果、方法和资料有机结合，从不同角度对考试现象全面、系统、多方位地探讨，这是现代科学发展的客观要求，也是深刻认识考试的特殊本质、强化考试理论基础的必需。

二、考试学与临近学科的关系

考试学因其内容宽大，相关学科甚多，但它与各相关学科的密切程度却不同一，关系的亲疏差别至为明显。在与其相关的学

①《马克思恩格斯全集》第3卷，人民出版社1972年版，第5页。

科中，尤以教育评价学和教育测量学最为临近，它们有着极为密切的关系。这主要表现在：

（1）考试、评价、测量均以人为对象；

（2）在教育系统中，三者都是根据一定的目的，通过一定内容，采取相应的形式、方法和手段，检验和判定受教育者在教育影响下身心发展的水平及各种教育活动的客观效果；

（3）在教育方面的功能十分接近，考试、测量、评价同是诊断教学问题、检验教学效果、评量教育质量的工具，同是督导教学改革的手段，同是实现教育目标的基本措施；

（4）学校考试、教育评价、教育测量各自的度量或评价标准，均以一定社会教育的性质、方针、政策所确立的课程、教学、教育目标为据；

（5）考试学、教育评价学和教育测量学的发展变化，均受制于社会发展和人的发展规律，为一定社会实践服务，是三门学科的共同归宿。

但是，它们之间的紧密联系，并未削弱考试学的独立性，相反加固了考试学的基础，使它与教育评价学和教育测量学之间的区别更为明晰。

教育评价的界说，迄今尚无统一固定表述。目前使用较多的定义是，"对教育目标以及对它们优缺点和价值评价的系统调查，为教育决策提供依据的过程。①"台湾学者李聪明则定义为："利用所有的评价技术评量教育所期的一切效果。"它的根本宗旨"是注重教育效果的价值观，依从新教育学及新心理学的指导，以现代教育的学力观为基准，分析教育应达到的目标，根据该目标来评定教育效果，以考查其实现的程度，作为再计划实

①中央教育科学研究所编：《教育文摘》报，1986年第24期第7版。

施的张本。”①其范围包括学业成绩、课程、教师、学校、各级教育质量等方面的评价。常用的方法为测试、调查、类比、评定等，“为教育决策提供依据”。

教育测量学是研究如何运用各种测试手段和方法，去测量受教育者在教育活动影响下所产生的客观效果。测量，并不是评价，史蒂芬斯曾把测量定义为：“就其广义来讲，测量是按照法则给事物指派数字”。作为测验的测量，它“只是着眼于客观数字的测定，其思想与价值观并无关联性”，②它为评价提供客观依据。

教育系统的考试，则是教学的重要环节，它本身具有教育性。考试对学业成绩、教学效果、教育质量的检测，既有测度的一面，又有评价的一面，故考试能左右教学方向、促进或抑制受教育者智能发展的功能。**教育系统的考试**的方法，通常采用的是书面质问、口头质问、操作、演示、设计等。其测试范围不涉及课程、教师和学校的评量。

由上可见，在教育系统中，考试、测量、评价三者的内容、范围、方法与功能，又各有特殊。固然，考试和测量同属评价的前提，亦可称为评价的基本方法或手段。没有考试或测量所获得的结果，评价便无判定依据，评价也就无法开展。但是，我们不能由此作出结论，认为考试和测量可归入评价之中。因为“一个大的事物，在其发展过程中，包含着许多的矛盾”。③这些矛盾各有其特殊性，如评价、测量和**教育系统的考试**，是教育这一大事物发展过程中的不同矛盾，有各自矛盾特殊性所决定的特殊本

①李聪明：《教育评价的理论与方法》，台湾幼狮文化事业出版公司1983年版，第3页、（前言）第2页。

②李聪明：《教育评价的理论与方法》，台湾幼狮文化事业出版公司1983年版，第5页。

③《毛泽东选集》（合订本），人民出版社1971年版，第286页。

质，可分别构成不同学科的研究对象，所以，教育测量学和教育评价学能成为教育科学的分支学科。

然而，考试学则又不同于教育测量学和教育评价学，它不属教育科学的分支学科，而是一门独立的学科。这是因为：

（1）在人类社会的母系统中，考试和教育是两个相互独立的子系统，虽联系甚紧，但互无统属关系；

（2）考试和教育是两种性质有别的社会活动：教育是施教者与受教者两种主体之间的社会交互活动，考试则是主试者与被试者两种主体之间的社会交互活动；

（3）教育是开发智力资源、劳动力再生产的手段，是培养人才的工具，考试是检测教育目标是否达到与促进教育实现培养目标、开发和合理利用社会人力资源的手段，是测度、甄别、任选人才的工具，是调节社会各部门人才结构的杠杆；

（4）在一个国家，考试制度和教育制度是平行制度，学校考试制度虽隶属于教育制度，但它只是国家考试制度的一个组成部分；

（5）教育系统的考试，仅是考试学研究内容的一个方面，考试学所要研究的是人类社会的考试现象，包括社会各系统的一切考试活动；

（6）尽管教育系统的考试与教育测量和教育评价有许多同一性，甚至考试在教育系统所涉范围还不如评价广泛，但教育评价和教育测量毕竟限于教育系统，其它社会领域人才素质和智能的测度、评价与甄选，不属它们的范围。

（7）教育评价和教育测量不具备考试所特有的社会功能：为社会各行各业甄别、选拔、任用人才提供客观依据；平衡社会所需人才的供求标准；调节社会各部门的人才结构；开发和利用社会人力资源等。

有鉴于此，我们认为，考试学与教育测量学和教育评价学有

相互交叉的地方，而又互有质的区别。考试学有其特定的研究对象、内容、任务、原则和方法，所涉范围远非教育能包罗。它如同教育学一样，是一门独立的科学，一门正在形成和发展中的新兴社会科学。

第二章

考试的定义、属性、产生及发展

历史是事物发生、发展的轨迹。
属性是事物内在矛盾的反映。

在本世纪70年代末至80年代初这短暂的十余年间，人们对考试的认识发生了历史性的转变，从前单以方法、手段看待考试的观点日趋消退，而视考试为一门科学的思想代之而起，并受到社会各界愈来愈密切的关注。那么，考试作为人类社会的一种特有现象，它的内涵、属性是什么？它们又是怎样产生、发展的呢？其演变又有何内在规律呢？迄今为止的考试究竟历经了哪些发展阶段？其发展的趋向又将如何？这正是本章所要讨论的问题。

第一节　考试的内涵

什么是考试？其内涵如何界定？这是考试研究应首先弄清的问题。

"考试"一词由"考"与"试"二字合成。在我国典籍中，对"考"、"试"二字，各有如下种种解释：

"考"字最早出现于甲骨文，与"老"同义。后随考试活动

的发展，人们逐渐赋予“考”字以新的含意。《广雅》中将考解释为“问也”，即询问之意。《尚书·尧典》有“三载考绩，三考，黜陟幽明，庶绩成熙”之说，依据董仲舒《考功名》的解释，所谓考绩，乃考其所绩也。又据班固《白虎通义》卷三《考黜篇》的解释，其意为三岁一考绩者，乃遐以适当时日，使三载有成始足以观之，也就是说，“三考而后定黜陟”，或谓以考核之法定升降。《礼记·燕义》又云：“国子存游卒，使之修德学道，春合诸学，秋合诸射，以考其义而进退之”，乃检查、询问学生德行和学业之进退。以上所云，考均为询问、检查、甄别人的德行、学识和劳绩之意。

“试”，《增韻》解释为“探也，较也”。其意是检验、比较。《新唐书·选举志》写道：“凡秀才，试方略策五道，以文理通粗为上上、上中、上下、中上，凡四等为第”。这里的“试”意指检验、比较秀才的才、德、学、识。可见，我国古代“考”与“试”两字的含意基本相同。

由“考”与“试”两字合成的“考试”一词，从我们已涉猎的资料看，它最早出现于汉代，如西汉董仲舒在《春秋繁露·考功名》中写道：“考试之法，合其爵禄，并其秩，积其日，陈其实，计功量罪，以多除少，以名定实，先内弟（第）之。”即是说，用考试之法检查、考核官吏的政绩，并按考试的结果计功量罪。

《三国志·魏书·王昶传》载：“昶陈治略五事，……其二欲用考试，考试犹准绳也，未有舍准绳而意正曲直，废黜陟而空论能否也。”王昶在陈述治国五项策略时，将考试列为第二项，提出用考试甄别官吏，检查政绩，辨别曲直，废黜空谈，从而揭示了考试的部分社会功能：考试是准绳，是评量和甄别国家官吏政绩的尺子。考试作为国策，其社会功能不仅仅是考核官吏，而且要用以选拔各种人才，督导教学。我国两汉、魏晋南北朝以前

的选士虽以荐举为主，但考试从未间断，《礼记·射义》云："天子以射选诸侯、卿、大夫、士。"又云："诸侯岁献贡士于天子，天子试之于射宫"。西周处于奴隶社会，高级官吏实行世袭制，但也须试射。初级官吏实行贡士制，所选举之士，天子加以考试。两汉官吏的选拔实行察举，也须考试，汉文帝15年下诏："诸侯王、公卿、郡守举贤能直言极谏者，上亲策之，傅纳以言。"①"对策者百余人，唯错（晁错）为高第，由是迁中大夫"。②魏晋南北朝实行"九品中正"选士制度，根据家世、状、品三项品评人物，以备政府随时录用，而官吏的任命须经考试，一般是孝廉考经书，秀才试策问。隋唐以后实行科举制，考试便成为甄别、选拔官吏和人才的主要手段。

上述历史典籍的有关记载表明，"考试"一词的含意，即是测度、考查、检验、甄别和评鉴。它的对象是人。其内容极为广泛，涉及人的德、学、才、识、体等方面。

到了近代，由于人们对考试的认识有所偏颇，反而缩小了考试的内涵，认为考试只是检测知识多寡和技能高低的一种手段或方法。我们认为，考试的概念具有广义和狭义之分。

一、广义考试与狭义考试

广义的考试，泛指人类社会一切测度和甄别人的身心各个方面之群体或个体差异的活动，如体能测验、性格测验、单元测验、期中考试、学期考试、学年考试、毕业考试、教学效果诊断考试、升学考试、招工考试、干部考核、业务考核、评比、鉴定、各种竞技比赛等等。凡人类社会具有测度、考查、检验、评鉴和甄别人的德、学、才、识、体等个别差异性质的活动，都属考试的范畴。

①《汉书·文帝纪》。

②《汉书·晁错传》。

狭义的考试，是由主试者根据一定社会的要求，在一定的场所，采取一定的方式方法，选择适当的内容，对应试者的德、学、才、识、体（诸方面或某方面）所进行的有组织、有目的地测度或甄别活动。因其性质、目的、内容、方法、手段的不同，狭义考试又分化为众多类型，现代社会使用较为普遍的有：合格考试，选拔考试，研究性考试；文字考试，非文字考试；个体考试，团体考试；深度考试，速度考试；论文式考试，标准化考试，变通式考试等，分别归属于社会和学校两大考试系统。

透过现象考察考试，它是主试者与被试者两种主体之间的一种交互活动。凡考试都有四个基本要素：主试、被试、测试内容与结果。考试以人为测试对象，其本质是测度或甄别人的身心个别差异的社会活动。

二、考试近义概念辨析

现代社会，考试形式众多，种类极繁，并有不少近义的概念，故有进行辨析的必要。我们认为，凡是测度和甄别人的德、学、才、识、体个别差异的社会活动，都属考试的范畴，但概念有准确、近义之别，内涵有大小、深浅之差。如，考查、考核、考察、考绩、考验；测验、测量、测试；评比、评价、竞赛、鉴定、检查，等等，都有考试的含义。但他们之中有的是考试的同义词，可以在一定条件下通用，如测验就是考试的同义词，期终考试，也可以说期终测验，但“高考”就不能说成“高测”。又如智力测验，不能称之为智力考试。我们说在某些条件下可以通用，也就是说，在某些条件下不能通用。原因在于：有的是习惯用法，有的是应用对象不同，有的是在含义上有细微的差别，有的只表示考试某一方面的意义，有的则是考试的引伸之意。辨析考试近义概念的目的，是为了正确理解考试的含义。现就考试的几个近义概念简要辨析于下：

考查：用一定的标准来检查、衡量人的某种行为或某种事件。若是对人的行为、活动进行考查（如对学生的学业成绩考查、对某负责人的政绩考查）它是考试的一种方式，属于考试的范畴。若是对一件事情进行考查，则不具考试的含义。

考核：考查、审定。如对干部职工的政治业务水平进行考核。它是考试方法的一种，属于考试的范畴。

考察：实地观察、调查。若对学生、职工学习和工作进行观察、调查，便是考试的一种方式，属于考试的范畴。若是对一项工程进行考察，则不含考试之意。

考绩：考查工作人员的成绩，是考试的一个侧面，属于考试的范畴。

考验：通过具体事件、活动或困难环境来检验一个人的品质、精神或意志力等，它是一种特殊方式的考试。如战争对一个战士而言，就是一种考验。

测量：运用仪器对事物进行客观描绘，并指派或确定有关数值。如大地测量、工程测量、气候测量、地震测量等等。后来发展为心理测验和教育测量。心理测验是从心理角度客观地对人的心理活动进行描述；教育测量是从教育角度客观地对教育效果进行描述。心理测验和教育测量虽具有某些考试的性质，但它只是给测试对象指派数字，不具有价值观。因此，心理测验和教育测量仅是考试的方法或手段。

测验：它由测量概念演变而来，如智力测验，学生成绩测验等。“考试”一词，在我国汉代已经出现（前已阐明），而“测验”一词在我国一直到本世纪初才出现。其含意是测定、检验，与考试是同义词。在一定条件下，它们可以相互替代，但在有些场合，两者又不能通用。这是因为，两者除有各自的习惯用法外，还在词的用途，含意上有细微的差异。在日常应用中，考试多用于比较正规的场合，而测验则用于一般的场合，如单元测

验，时事测验。在西方，将测验定义为"经过标准化的测量"，从这点意义上讲，测验是客观测量，而考试包括主观考试和客观测量。这就是它们的原始差异所在。

测试：是精度的测量。应用于人的素质与智能的测度时，它属于考试范畴；若用于物的测度，它则属于测量学的范畴。

评价：其一，评定价值高低。即通过分析、比较、对某人给以正确的估价，这属考试的一种方法。其二，评定的价值，即对某人评定的结果，或对此种结果的表述，或对测量、考试所获结果的看法和结论。应当指出，评价既可用于人，也可用于事和物，只有用于人时才有考试的意义。

鉴定：即对人的优缺点作出鉴别和评定，是考试的方法之一。

考试的近义概念很多，只能择要而述，不可一一穷尽。上述概念究竟如何准确界定，目前仍有不同看法，有待深入探讨。

第二节　考试的基本属性

考试的属性，也就是它的性质和特征。研究考试的属性，有助于认识和掌握考试的规律，为考试实践提供可靠的理论根据。毛泽东同志曾指出："不论做什么事，不懂得那件事的情形，它的性质、它和它以外的事情的关联，就不知道那件事的规律，就不知道如何去做，就不能做好那件事"。① 从事考试工作，必须认识和掌握考试的规律，才能使考试工作建立在科学的基础上。

所谓考试规律，指考试发生发展的内在联系及其必然趋势。考试规律是不以人们的主观意志为转移的客观存在，它决定考试

①《毛泽东选集》（袖珍本），人民出版社1969年版，第155页。

发展的基本过程和方向，具有普遍性、重复性的特点。

如果说事物内部矛盾的外化是事物的属性，那么，考试内部矛盾的反映于外，就是考试的性质和特点。考试的性质是由社会政治经济形态所决定的，而其特点则是考试自身发展规律的派生现象。考试的基本属性可归纳为四个方面：

一、永恒性与时代性

人类社会发展和延续，必须具备两个充足条件：其一，人类必须不断地创造物质财富和精神财富，以满足自身生活和社会扩大再生产的需要；其二，人类需有自身的再生产，繁衍后代。这是人类社会发展和延续的根本法则。而人类创造物质财富和精神财富的唯一途径，又在于分工合作的劳动。它不仅创造了社会，同时也创造了人类本身。人是不能脱离社会而生存的。也就是说，人类劳动的分工合作和生活的相互依赖，既是人类生存的基础，也是每一个社会成员生存的基础。尽管在不同的社会中，劳动分工合作的形式不尽相同，从事劳动的性质也不完全一样，但劳动是人类与其他动物本质区别所在，为人类创造物质财富和精神财富所必需，也是人的个体自身需要。考试是甄别和测度人的个别差异的活动，为社会分工提供依据，也为个人取得相应的工作职位或学习机会提供条件。因此，只要社会存在分工，考试就必然存在。从这种意义上讲，考试这种社会现象具有永恒性。

再者，人类自身的再生产，需要教育对年轻一代进行培养。而教育伴随人类社会的始终，它是一个永恒的范畴。考试和教育是一对孪生子，教育的本质是培养人，考试的本质则是甄别和测度人，它必然贯穿教育的始终，两者相辅相成。既然教育是人类社会永恒的范畴，那么，考试也同样是人类社会永恒的范畴。

考试的永恒性是考试的基本属性，但这并不等于考试性质固定不变。随着社会的发展、科学技术的进步和分工的变革，考试

为适应不同时代的人才需求，它始终在变化性质、更新内容和改变形式，从而表现出浓厚的时代色彩。比如，原始社会的民族或部落中，16岁或18岁的青年，只有经过冠礼或命名仪式，才能成为民族或部落的正式成员，享有民族或部落成员的权利和义务。这种冠礼或命名仪式，即是一种原始而又严肃的考试。奴隶社会，古代希腊斯巴达农业贵族为了训练他们的子弟成为武士，组织青少年举行夜间捕杀奴隶的“实习”，以考验他们的所谓“勇敢精神”，实际上是一种为检验阶级意志而进行的残酷的考试。我国历史上的科举考试，近现代英国的文官考试，美国高等学校招收新生所进行的学能考试，苏联和东欧各社会主义国家的高等学校入学考试，我国高等学校统一招生考试、科技工程人员业务职称考试等等。这些不同时代、不同社会的考试，为其所处社会或时代的性质、生产力水平而决定，在性质、内容和形式上都有不同。或者说，为社会的政治经济和科学技术的水平所决定，在不同的时代，社会对人的素质和智能有其自已的要求与标准。考试的性质、内容和形式随其变化而变化，从而赋予考试的时代性。但这并不排斥考试本身的规律和对历史上考试优秀遗产的继承，也不等于对不同社会或时代的考试先进经验的排斥。

二、社会性与独立性

考试是以人的个体或团体为客体的社会活动，其社会性是客观存在的。考试的社会性有两种含义：

其一，在无阶级的社会中，考试的社会性体现为全民性。其对象是社会全体成员，目的在于为全社会创造物质文明和精神文明服务。其职能是对社会成员的素质与智能进行甄别和测度，为各类社会人员的选拔和评价提供依据，以求社会人力资源得到合理的开发和利用，使社会不同素质与智能的人，都可各尽其才，成为推动社会前进的一份力量。

其二，在阶级社会中，考试的社会性体现为阶级性。其对象是社会中的少数人，主旨是为统治阶级服务。其职能在于甄选统治阶级所需的人才。如中国历史上的科举考试，就是为封建王朝选拔官吏。英国的文官考试，同样是为巩固资产阶级政权而设置的；按照资产阶级的意志选拔官吏，是文官考试的基本宗旨。就上述意义讲，考试是一种工具，谁掌握考试的支配权，考试就为谁服务。

但是，考试又有其独立性。它的发生、发展有其自身的规律，不为阶级意志所左右；它的功能的正确发挥，必须遵循其原则与规范，任何随意性都会导致功能走向反面。也就是说，考试的社会性虽然随社会性质的变化而有所区别，但其内在规律和固有功能，却不随人们的意志而改变。

三、主观性与客观性

衡量考试成功与否的标准是：主试者欲测的内容是否与考试所测的结果一致；考试的结果与应试者被测方面所具有的实际水平是否一致。如果考试内容、考试结果、应试者被测方面的真实水平三者一致，这说明考试是成功的，达到了主观性和客观性在考试中的辩证统一。这是因为，考试的目的是根据社会的需要和应试者的身心特征由主试者确定的，社会的需求和应试者的身心特征是客观的，而主试者对这两者的认识和目的的确定则是主观的；测试结果是主试与被试两种主体交互作用的产物，它是客观的，但对测试结果的解释却又渗入主观的因素。所以，考试目的、内容的确立，结果的形成和结果的评价，都无不含有主客观两种因素。

四、有限性和无限性

有限和无限是一对哲学范畴。恩格斯在《反杜林论》中指

出："人的思维是至上的，同样又是不至上的，它的认识能力是无限的，同样又是有限的。按它的本性、使命、可能和历史的终极目的来说，是至上的和无限的；按它的个别实现和每次的现实来说，又是不至上的和有限的。"①人类的考试也具有有限和无限的基本属性。

考试的有限性，是就考试的对象和内容及其可测范围而言的。也就是说，任何一种考试，它的对象都只是社会成员的一部分，其内容只是被测方面的主要成分，即人的德、学、才、识、体某一方面或几方面，而不能是全部社会成员及人的个别差异的所有内容。如中国古代的科举考试和英国现代的文官考试各自的测试对象都有明确的限制。科举考试要求考生投保，保证出身、职业、品德等符合政府规定的要求；文官考试也是如此，考生必须达政府规定的学历等等。考试的内容有限定的范围，录取的人数有一定的限额。这都反映了考试的有限性。

考试的无限性，是指考试的影响或教育意义。一种考试的实施及其结果的利用，往往影响同类人员的学习、工作和奋斗目标，乃至后继人才智能发展的方向，有时还会影响一代甚至多少代人。所以，良好的考试能给同辈、后世者以积极影响，进而产生良好的社会效益。相反，若考试缺乏科学性，不能反映社会发展的客观要求，也必将为同类或后继人才导偏方向，甚至埋没、断送人才，成为社会前进的障碍。事实也是如此，历时1300年的科举考试，在历史上既起过促进作用，也起过阻碍作用；既选拔过优秀人才，也扼杀过无数人才，其影响远远超出它存在的时间和中国的地域，直到今天，英美的文官考试仍然还留有科举考试的烙印。所以说，考试又是无限的。

总之，考试的基本属性是考试内在矛盾的客观反映。在考试实施过程中，只有把握住这些规律性的特征，才能正确发挥考试

①《马克思恩格斯选集》第3卷，人民出版社1972年版，第126页。

的社会功能，对社会的进步起到积极的推动作用。

第三节　考试的演变与发展

人类的考试活动，纵贯古今，横涉中外，欲将考试产生、发展、演变的历史轨迹清楚地勾画出来，不仅需有长时期的收集、整理和考证，而且要用较多的文字进行叙述，故拟在中外考试史中再细致探讨。本节只对考试发展、演变的基本线索作概要阐述，以揭示考试现象发展、演变的内在规律。

一、考试的萌芽

考试产生于何时？目前仍是中外考试研究工作者竭力探究的重要课题之一，比较一致的看法是：有教育就有了考试，考试与教育是一对孪生子，它们因社会生产和社会生活的客观需求而同时问世。具体说，考试与教育的产生需具备下列条件：人类有了制造并采用工具从事物质生产的社会实践活动；有了维系社会成员生存与发展的基本生产技能和生活经验；为使生产劳动得以持续，人类及社会生命延绵不绝，社会提出了给青年一代生产技能、生活经验传授，以及分辨新生一代掌握技能熟练程度、接受生活经验快慢的要求；人类产生了语言，出现了思维活动，有了清楚的意识和一定的抽象能力及推理能力，劳动与生活成为有意识的活动。这些条件大约到了人类蒙昧时代的后期基本具备。所以，考试与教育可推定萌芽于原始社会的蒙昧时代。但由于具体情形缺乏文字记载，因此不得不依靠考古学家提供的有关资料和现今世界上仍处于原始形态的部落的教育和考试活动，来讨论原始社会萌芽状态的考试。

（一）原始社会考试的性质与目的

原始社会是公有制的社会，共同劳动、共同享用，没有剥削、没有压迫，氏族或部落的一切大事，都由全体社会成员以民主方式处理，青年一代的抚养教育是氏族、部落的任务，对青年一代的考核也是氏族、部落的职责。这一时期的教育和考试，都在生产劳动和生活的实践过程中进行。无论是教育还是考试，都是全民性的。

现阶段，居住在中亚细亚地区的原始部族，其生活境况尚处于人类蒙昧时代，他们仍是一个个无性别分工的集体。每个集体包括两群，一群是猎捕大兽的成年男女，另一群是负责饲养动物、守护家园的老人和儿童。在这一群体中，老人对儿童和少年进行护理和教育，并在教育中进行模仿快慢、动作熟练与否、适应能力强弱、力量大小等方面的考核。此种教育和考试的情景，基本再现了考古学家和史学家所描述的人类蒙昧时代考试与教育的状况。当历史发展到母系氏族社会后，男女两性分工便有所不同。在现今中亚仍处于这一历史阶段的氏族中，男女在8岁以前生活在一起，由妇女负责护理。8岁以后男童由成年男子指导，学习狩猎等事，女童由妇女教以饲养等事，并通过实际活动检测学习的成效。

在父系氏族社会里，教育与考试则由以血缘关系构成的氏族或部落中的长者主持。如美洲的印第安人，男童和女童及女青年，一般是由氏族或部落的成年妇女教导与考试。当施行某些严格训练与考试时，则由生母的兄弟来协助，祖父和外祖父也来帮忙，主要检验她们的生产技能和氏族的传统风俗与成训。至于男青年，除考核生产技能和氏族风俗及成训外，还要进行军事及民族历史知识等方面的考核。

由于原始社会青年一代是公养，所以，这一时期的所有青少年都享有平等教育权利，考试对所有的青少年也是一律平等。考

试的目的，在于使青年一代成为合格的氏族或部落成员。

（二）原始社会考试的内容与方法

1．原始社会考试的内容。

原始社会生产力低下，新生一代的抚养和教育依靠群体。由于受教育者学习的内容是社会生活和生产劳动的实际知识与技能，所以，这一历史时期的考试，主要包括下述几方面的内容：

第一，氏族、部落的成训、禁戒，部落的图腾，本氏族或部落的历史传说，习惯风俗和地理环境等。

第二，有关采集、狩猎、制造工具、营造住所、播种、收割、饲养等方面的知识、经验和技术。

第三，宗教礼仪，音乐、舞蹈、绘画、体育方面的知识，部落或氏族的大事，英雄人物的业绩，结绳、贝壳、彩色石子的含义等。

第四，军事知识，军事技能与征战本领，勇敢精神与意志力等。

2．原始社会考试的方法。

原始社会的教育与考试是在社会生活与生产的实践过程中进行的，考试方法以实践为主，同时辅之以口头陈述与实际演示。

原始社会的儿童考试，是通过游戏与模仿成人行为的方法进行。如非洲儿童通过模仿成人设陷阱猎兽的游戏。亚马逊河谷的儿童用仿制陶器的游戏，来接受教育护理者的评价。游戏检测中的获胜者，可得到教育护理人的口头表扬或鼓励性奖品。而美洲的印第安人，则是在天寒季节将儿童集聚一起，由年长者讲述民族传统和历史故事，次日以复述的方法对儿童进行考试。

原始社会对青年的考试，最庄严的是氏族或部落举行的青年礼，又称命名礼或冠礼。美洲的印第安人和非洲、澳洲的一些部落，凡成年的男女，都要接受多方面的考试，成绩合格者才能获得氏族成员的资格。尽管因部落不同，考试的具体方法和要求不

尽一致，但都是以严肃而庄重的青年礼之类的方式进行的。青年礼由部落中德高望重的长者主持，对适龄青年身体、道德品质和劳动能力，以及部落的历史法规与戒律等，进行严肃的考试，或令其单独居住一段时间，或让其忍受饥寒与体罚，或要求他们从事猎兽或其他繁重的劳动等等。参加仪式的青年，在考前有较长的准备期或训练期。如印第安的阿帕契部落就规定，参加典礼的女青年要有一年的准备。有的部落青年，在参加青年礼以前须拜人为师，掌握一些军事知识与技能。有的部落给长期准备的青年特设营地于森林之内，男青年训练2～8年，女青年至少训练一年。在训练期间不得离开营地回家，营地由部落高级首领负责监督和严格训练，并传授有关部落生活、生产等各方面的知识和技能。凡传授、训练的内容，都要进行严格的考试。

原始社会氏族、部落的首领由民主选举产生，同时氏族、部落的议事会有撤换酋长和军事首领，以及其余的"信仰守护人"的权力，一切成年男女享有平等表决权。这种选举实质上是一种考试，也是现代人才考核、选拔的先声。我国《尚书》、《史记》记述有尧舜时期的选举情形，尧的哥哥挚曾是部落联盟的首领，由于他为人"不善"而被罢免，由尧接替；尧晚年询问"四岳"（四个部落酋长），有谁可继承他的职位，四岳推荐舜，尧表示要对舜考验一番。舜被推荐做尧的继承人以后，尧"乃以二女妻舜以观其内，使九男与处以观其外。……于是尧乃试舜五典百官，皆治"，"舜得举用事二十年，而尧使摄政，摄政八年而尧崩。"①这说明原始社会对领袖的甄别、选择是严格认真的，而且有一套考核制度保证。同时也说明，在原始社会时期不仅有了考试的现象，而且有了有意识、有组织的考试实践活动，并有多种测试内容和方法，以及人类最原始的考试制度。尽管这些考试还不是完全意义上的考试，但它实质上具有考试的性质，应是

①《史记·五帝本纪》。

人类考试的先声，也是当今考试之源，在人类社会的考试发展史上有其特殊的地位。

二、考试的发展和分化

随着原始社会的解体，人类社会进入了古代文明时代——奴隶社会、封建社会。考试也由萌芽状态的民约、礼仪式考试，逐步发展和分化为成文的法规考试和不成文的民间技艺考试。

成文的法规考试，是统治阶级——奴隶主、封建地主阶级施行政治统治、经济剥削和精神奴役的工具。考试的对象是统治集团的子女和剥削阶级的后代。测试的内容是统治集团规定的经典知识和治国之术。目的在于选拔国家官吏，培养教育维护本阶级利益的人才。成文的法规考试后来发展为社会考试和学校考试两大系统。社会考试的主要职能：一是选拔人才；二是检测学校教育质量。学校考试的主要职能：一是检测学生学业成绩；二是为改革教学服务。

不成文的民间技艺考试，是民间流传的没有明文规定的依地区或行业习俗而形成的考试，其目的是检测生徒掌握技艺的熟练与准确程度。其对象是奴隶、下层自由民、农民及手工业者。测试的内容是本行业的艺技、戒训及戒律。其中最典型的是手工业者的师徒授业考试。

考试的发展和分化，是社会进步所带来的必然结果。原因在于社会生产力的发展，私有制的出现，阶级的产生，国家的形成，社会政治、经济、文化科学的发展对考试产生了新的需求，促进了考试的自身演进。考试出现上述发展分化的条件有二：

其一，当原始社会进入阶级社会后，社会生产力有了较大发展，科学技术明显提高，奴隶、农民创造了较之原始社会更丰富的物质财富和精神财富，文字不仅产生、形成，并为社会广泛使用，书写手段不断更新，从而为考试的发展和分化提供了必要的

技术和物质条件。

其二，考试是客观存在的社会现象，有其自身发展的规律，只有当人们认识其规律并使其为社会服务时，才能产生社会价值。人类进入阶级社会后，由于政治、经济对考试有了新的需求。人们通过实践对考试的规律和功能的认识进一步深化，能够根据本阶级的要求，有意识有目的地甄别和测度人的德、学、才、识、体个别差异，以实现有效培养、选拔、任用人才的目的。

（一）考试的发展

考试的发展决定于社会发展的需要、社会提供的物质条件和人们对考试的认识水平。

在漫长的奴隶社会和封建社会，考试发展是不平衡的。在中国，考试已由幼芽长成为枝叶繁茂的参天大树，并形成了一个完整的考试体系，渗透于社会肌体，服务于社会的物质与精神文明；而在西方，这一时期的考试却处于幼芽状态，仅初具雏型。鉴于我国这一时期的考试在人类考试发展中具有典型性，便以我国奴隶社会和封建社会考试发展、演变的史实为例（当然也要涉及西方某些典型的考试史料）进一步讨论原始社会后考试发展分化的轨迹与成因。

1．考试制度的发展。

考试制度是成文法规考试的重要特点，也是考试发展趋向成熟的主要标志。原始社会考试处于萌芽状态，考试仅随习俗而成。到了奴隶社会及封建社会时期，考试制度便成为国家的重要制度之一，如我国西周以后选拔官吏的制度有：贡士制、养士制、军功制、察举和征辟制、九品中正制和科举制等。在职官吏的考绩制度有：述职、巡狩制、上计、刺察制、司考、校考制、磨勘、历纸制、考功历制、考满、考察制、汇考制等。在学校考试制度方面，西周已有学校考试制度的记载，《礼记·学记》

云："比年入学，中年校考。"《周礼·地官司徒》亦云："乡大夫之职，各掌其乡之教政禁令。正月之吉……以考其德行，察其道艺。……三年则大比，考其德行道艺，而兴贤者能者。"还有"春合诸学，秋合诸射。"古代太学已有一套比较完整的考试制度。西汉每岁一试，"设科射策"，试而优则仕，不及格者令其退学。唐代有"旬试"、"岁试"和"毕业试"，以成绩优劣分别给予升降。宋代王安石改革太学制度，实行三舍法，择优授官。

13世纪以后，欧洲城市商业和手工业有了新的发展，手工业者建立行会组织，青少年投师学艺，有2～3周试用期，考查投师者是否适合学艺，品行上有无劣迹，合乎要求者方可招收为徒弟，登记入行会名册。徒弟出师、工匠升为师傅，也须经过行会考试。

以上事实说明，这一时期的考试已渗透于社会政治、经济、教育领域，各种考试制度日趋完善，从中国官吏的选拔考试的发展过程，可以看出考试制度发展演变过程的概貌。

早在2000多年前的西周时期，尹、史、卿、大夫等高级官吏实行世袭制，只在下层贵族中选拔低级官吏和武士时，才实行贡士制。据《周礼》记载：乡大夫每年正月对本乡人士，"考其德行，察其道艺。……三年则大比（考试），考其德行道艺，而兴贤者能者（选拔德贤才高之人）。乡老及乡大夫帅其吏，与其众寡，以礼礼宾之，厥明（厥，其也。其宾明日也）乡老及乡大夫群吏，献贤能之书于王。"①诸侯贡士，天子试之于射宫。"②西周的贡士制是推荐与考试相结合的制度，选拔的标准是贤与能，贤是奴隶社会的德行，"德"即六德："知、仁、圣、义、中、和。"③"行"即六行："孝、友、睦、姻、任、恤。"④能是奴隶社会的道艺，也就是"六艺：礼、乐、射、御、书、数。"⑤考试是每年一小考，三年一大考。乡选之士，献于朝

①②③④⑤《周礼·地官司徒》。

廷，天子再进行复试。西周初期由于征伐天下和镇压商族叛乱及奴隶起义的需要，贡士制选拔的目的是录取低级官吏和武士，因此非常重视射御等军事技术的考试，故有天子试贡士于射宫之说。

贡士制突出的局限性在于，录取官阶仅限于低级官吏，不过是世袭制的一种补充形式。考试的内容和方法过于抽象，施测不甚规范。但相对原始社会的考试来说，贡士制却有长足的进步，它开创了通过考试选拔官吏的新纪元。以后春秋战国的荐举和养士制，秦代的军功制度，两汉的察举征辟制，魏晋南北朝的九品中正制，隋唐至明清的科举制，都是它的继承与发展。荐举制和养士制虽没有直接考试的形式，但询问、陈述含有考试的因素；军功制按军功授予官爵，是实践性的考试；察举制和九品中正制是推荐制，但也需经过考试后，才能授予官职；科举制则是以考试成绩优劣为标准，甄选任用官员，它是由贡举制发展而来的最完善的考试制度。但因历史上统治阶级的政治偏见和科学技术等多方面的限制，以上各种考选制度都有其弊端和先天不足之处。所以，随着社会政治、经济和文化科学的发展，以及社会对人才培养、选拔要求的不断变化，人们对历史上的考试制度有继承，更有创新和发展。

2．考试诸因素的发展。

考试是一项系统工程，它由各个因素组合而成，诸如：考试内容、分类、设科、方法及组织管理等。考试诸因素的发展、完善和科学地集合，是考试自身发展的重要原因。

（1）考试内容的发展。考试内容的发展决定于各个历史时期社会政治经济的需要，它表现为各个阶级对人才德、学、才、识、体各方面的要求，这种要求又是与当时的文化科学知识发展的水平相关联的。比如，西周初期，由于征伐天下以军事为主，需要勇猛善战的大批武士。因此，选拔贡士重在试射，严格检测

应试者射箭的技术与体力。到西周后期，由于天子治天下以礼乐为国策，选拔人才的标准及考试的内容也随之而变。如天子测试贡士射箭所注重的是，贡士射箭的体态是否合乎礼的规范，射箭动作的节奏是否合乎乐的要求。又如隋唐以后的科举考试，虽都是以儒家经典为考试的中心内容，但各代的考试内容又有程度不同的变化。在唐代的科举考试中，《论语》、《孝经》是各科必考科目，答卷仅以此为依据进行论述，可以发表自己的独立见解。而明清科举考试，试题则全出自"四书""五经"，答卷必须以此及朱熹《四书章句集注》为标准，迫使考生套用古人语气，代圣贤立言，不允许有自己的独立见解。将考试内容公式化、僵死化，是当时统治阶级政治的需要。明清已处于封建社会晚期，社会危机四伏，阶级矛盾尖锐，没落的地主阶级企图借理学的思想窒息进步思想意识的发展，以维护其腐朽统治。

（2）考试分类设科与分级筛选。西周贡士选拔尚未分科，但已有分级升等的选士、俊士、造士、进士的雏形。隋唐以后，考试发展分为常科与特科（制科）。唐代常科设有秀才、明经、进士、明法、明算诸史、明字等科，定期举行考试；还有不定期举行考试的俊士、一史、三史、三礼、三传、道举、童子及制举等科。特科有76种之多，如：时务、师范、著述、孝廉、茂才、博学宏辞、贤良方正、直言极谏、将帅科等等，由皇帝下诏，按需要设置、不定期举行考试。

唐代选拔考试分两级三个步骤进行，第一级预备考试，分为学校考试和社会考试。考试合格者，前者称之为生徒，后者称为贡士。第二级为国家级考试，首先是省试，参加者是生徒和贡士，考试合格者就取得了做官的资格。其次是吏部的"铨选"考试，合格者按考试成绩授予官职。

科举考试发展到明代，其正式考试分为三级进行，即院试、乡试、会试和殿试。在院试前须经过县试和府试二级预备考试。

考试分类设科，分级筛选，是考试的重大发展。分类设科是为了选拔各类专门人才，考试内容既有公共必考内容，又有各专业相关内容，故有利于人才选拔。分级筛选使考试的竞争更强，选拔更准确。这种选才方法具有一定的科学性，我国现行的高校入学考试基本上是沿袭此法。

（3）考试方法发展多样化。考试方法也是随着文化科学技术的发展、社会物质条件的提高而不断变化发展的。西周的考试方法，仅有口试、实际操作与演示，如询问与试射。到了唐代，由于笔墨纸等书写工具的广泛应用，考试方法除了口试、演示以外，还有帖经、墨义、策问、诗赋等方法。考试方法多样化，对人的素质和智能测试更加全面，既能测试考生的记忆能力，也能测试考生思维表达等多方面的能力，所以有利于人才的选拔。

（4）考试的组织与管理。西周的贡士制、两汉的察举制、魏晋南北朝的九品中正制选拔之士，先由地方推荐，然后参加考试复查。当时选士以推荐为主，考试处于次要地位。隋唐以后，科举制度以考试成绩选士，考生不经过推荐就可自己到官府报名，从而扩大了考试对象的范围，拓宽了生源渠道。

考试的主考官及同考官由朝廷任命，甚至皇帝亲自担任。到明清以后，任命主考官及同考官须经过考试。委任后的乡试主考官根据路程的远近限期到任，实行“锁院制”，自宣布之日起，不得探亲访友，不准携带家属及过多随员。会试在京城举行，被任命的会试主考官一经宣布，就必须立即到院上任，不得回家或投书联络。

两汉、魏晋南北朝时期的察举和九品中正制，选士以推荐为主，考试处于辅助地位，考试的组织管理比较简单。隋唐以后的科举考试，是以考试的成绩高低为录取标准。因此，科举考试一开始就非常严格，但组织管理工作尚不十分严密。唐代考生入试场虽要搜身，但试卷不糊名，考生在试场对试卷上不明确的问题

还可以提问，请主考官解释。评阅试卷仅在卷上批上“通”或“不”一字即可。宋代以后考试组织管理工作不仅非常严密，而且非常完善。清代尤甚。其试卷糊名密封，并盖有官府印记；考生答卷不得述说自家身世门第，也不能要求主考官解释试卷不明白的问题。会试、乡试设内外帘官分管监考、试卷、考场工作。内帘官主管阅卷、监试、管理试卷等。外帘官主管监试、考场事务。外帘官不得进入内帘官所在地，内帘官不得走出贡院，内外帘官不得来往。考生答卷由誊录房用朱笔重抄、对读后，将誊写的朱卷交同考官评阅，按○、△、╲、一、×五个等级，由八人轮流传阅评卷，以“○”多为优等。考生原卷——墨卷由外帘官保存，以便磨勘。

考试的发展，从制度到组织管理，均是由简单到复杂，由单科到多科，由个体测试到群体测试，由口试、操演发展到口试、笔试、操演多种方法的测试，是一个由不完善逐步达到完善的过程，也是一个除旧布新不断变革的过程，不断适应社会发展需求的过程。

（二）考试的分化

奴隶社会、封建社会都是农业社会，以农业为主体，手工业、商业、运输业均处于从属地位。这一时期，奴隶主、封建地主阶级掌握着政治经济大权。阶级的对立、封闭的自然经济和体脑分离使考试逐步分化为官府成文的法规考试和民间不成文的技艺考试，成为目的、性质不同，内容、方法各异的两大考试体系。

1．考试性质和目的的差别。

官府考试是从原始社会全民性萌芽状态的考试蜕变成为统治阶级服务的阶级考试。考试的对象是统治阶级的成员，但在内部又有严格的等级限制，如西周贡士制的考试对象仅是低级的贵族，选拔低级官吏与武士，高级官员却采用世袭制，由高级贵族

就任。魏晋南北朝的九品中正制，门第是品评的重要条件。元代科举考试对蒙古人、色目人、汉人、南人的要求各不相同，阶级限制和民族歧视十分严重。隋唐以后的科举考试，虽然考试对象的范围有所扩大，但主要限于剥削阶级内部，劳动人民因没有参加考试的条件，也就剥夺了他们参加考试的权利。总之，官府考试是为统治阶级服务的，具有浓厚的阶级色彩。民间考试是原始社会萌芽状态考试的继续，仍然保留其全民的性质，不受阶级出身门第的限制。手工业收徒授业不管是农民子弟，还是自由民子弟，甚至破产的剥削阶级子弟，只要通过某种考试，尊师守规，都可接受。中国封建社会民间的私塾蒙学，招收蒙童，地主、商人、手工业、农民及皂吏的子弟，只要备以束脩都可入学，参加塾师学业考试。

官府考试不管是社会的选拔考试，还是学校考试，都是为选拔和培养统治阶级御用人才；考生的目的在出仕做官。贡士制、察举制、九品中正制、科举制是如此，汉代的太学考试、宋代王安石在太学所实施的三舍法也是如此。而民间考试则是检验学徒专业技艺掌握的程度，出师考试则是为了取得社会承认学徒谋生合法地位，也标志学徒掌握了某种专业的技艺，有了独立从事某种事业的能力如从医的学徒出师后，即可挂牌行医。

2．考试内容与方法的差异

考试的内容为考试的目的所制约。官府考试的内容是统治阶级的经典知识及治国之术。中国的科举考试是以儒家的“四书”“五经”为内容；中世纪的欧洲教育为教会所垄断，不论哪类学校，圣经都是必考的内容。

中国封建社会考试的内容侧重于伦理知识，而欧洲封建社会的考试内容侧重于宗教神学，其最大的弊端是鄙薄生产知识和科学技术。而民间考试则以生产经验和专业技艺为主要内容，以及行会戒律与戒训等。显然前者考试的内容是书本知识，后者的考试内容则是实践性的专业技能。产生这种现象的原因，在于两种

考试的性质不同。

考试的方式和方法与考试目的、内容紧密关联。官府考试的方式是由个体向团体演变，贡士的试射、察举的策问都是个体考试，科举考试则发展为团体考试。民间考试仍然保持个体考试的方式，不管是学徒出师考试，还是私塾的学业考试都是如此。

官府考试方法的发展较民间考试发展演变快，对现代考试有深刻的影响，使考试方法发展多样化。如科举考试的口试有口义、策问；笔试有墨义、帖经、诗赋等。既有客观性的固定应答，又有主观性自由应答，还有操演性考试方法，如武举的武术操演，艺术学科的书法、绘画实践。考试方法的多样化对测试人才是有益的，今天的考试方法正是上述方法的继承与发展。

民间考试则以实践性操演为主，口试为辅，笔试次之。

考试的分化，既是人类分化为不同阶级的结果，又是农业社会自然经济的产物，也是考试自身发展的必经阶段。随着商品经济的发展和现代大生产的兴起，以及现代科学技术的迅猛发展和日趋综合化，考试的发展又由分化逐步走向统一，产生新的飞跃。

三、考试的飞跃与统一

考试萌芽于原始社会，发展、分化于奴隶社会和封建社会，飞跃、统一于现代社会——资本主义社会和社会主义社会，这是生产力发展和生产关系的变更在各个历史时期对考试提出不同要求的必然结果。在现代社会里，没有文化的劳动者不可能适应现代大生产的要求。很难设想，一个不会认图、不会计算的人，能承担钳工、铸造工、车工、装配工等机器行业的生产任务；一个不懂图表、仪器和化学反应的劳动者，能在化工企业的生产中出色完成任务，冶炼、纺织、交通航运部门也概不例外。现阶段，不管是工业、农业，还是服务行业、都要求劳动者有相应的文化知识素

养和科学技术能力。同时，科学在发展，技术在革新，大生产的基础是革命的，它将不断地促使人们从一个生产部门转向另一生产部门，从一个行业转到另一行业。要适应部门的转换、行业的变更以及科学技术的发展，没有宽厚的知识和技术能力是不可能的。不管是资本主义社会，还是社会主义社会都是如此。当然，在社会制度不同的国度里，其政治信仰、社会道德观念和人生观观等是有着本质区别的。如社会主义国家以实现共产主义为奋斗目标，它要求人们具有共产主义人生观，坚信社会主义制度，有理想、有道德、有文化、有纪律。而资本主义国家，则要求人们拥护资本主义制度，维护资产阶级利益，遵守资产阶级的法律制度。由此，便从生产方式和社会政治制度需求出发，对人才的培养和选拔提出了新的标准。考试为完成社会赋予的使命也由发展分化阶段过渡到飞跃统一阶段，即从旧的手工式考试演变为现代电脑化手段的考试，由主观考试发展为客观考试，并逐步形成论文考试与标准化考试相结合的新型考试模式。考试的这一演变过程。在东西方许多国家经过了数百年时间，下面对其演变脉络进一步探讨。

中国是考试的发源地，处于考试发展、分化后期的科举考试，在我国延续1300年，它不仅对我国现代的考试影响极深，而且成为欧、美、亚许多国家近代文官考试仿效的模式。最早向西方介绍中国科举考试内容与方法的是葡萄牙修士胡安·贡萨雷斯·德万多萨。他在1583年出版《伟大的中国》一书中，用一章的篇幅全面叙述科举考试。英国在1570～1870年间出版的《文官考试制度》、《中国札记》，《中国的历史与现状》等书中，极力称赞中国的科举考试，并主张英国政府仿行。当时英国驻华使节也向本国政府汇报过科举考试的内容。英国《绅士杂志》、《伦敦杂志》等著名刊物，先后撰文介绍中国录用官员的程序和方法，提出了英国以中国文官考试方法选拔官员的主张。到17

世纪，中国的文官考试制度终为英国效仿，并成为英国后来一种定型的考选制度。1853年，英国议会任命马莱组织委员会调查东印度公司的职员任用制度。他在调查报告中提出，通过考试择优录用东印度公司职员，并在东印度公司任用职员时付诸实施。这是英国通过考试择优录用官员的开始。与此同时，英国财政大臣格莱斯顿委托查理·屈维廉调查政府人事行政的状况，拟成《关于建立英国常任文官制度的报告》，提出凡欲担任政府事务官者，必须经考试录用。1855年5月上任的帕麦斯顿政府颁布了《关于录用王国政府文官的枢密院命令》，英国文官（事务官）必须通过考试录用。英国通过考试选拔文官的制度由此而形成。

在美国，1851年国会众议院部分议员提出建立官员分级考试制度。1853年和1855年国会先后通过两个法案，规定部分官员录用必须经过考试。但由于法案不具体，不久便流于形式。但美国学习、借鉴中国考试制度的思潮并未因此消失，而是更加高涨。1856年国会授权总统制定考选文官的规则。1866年10月北京同文馆馆长马丁，于美国波士顿东方学会作了“中国的竞争考试”的报告，向美国人介绍科举制。1867年10月《北美评论》杂志撰文称誉中国的考试取仕。1868年国会参议员托马斯·詹科斯以“节约联合委员会”名义向国会提出的报告中，专辟一章论述中国的文官制度，报告的这部分内容引起议员们的注意和兴趣。1887年海斯总统命令海关官员和税务官员经考试录用，但未能真正实行。而在美国首先开始用考试录用官员的是当时的内政部，只是未加推广。1883年国会通过了彭德尔顿文官法，规定政府事务官必须经过公开考试择优录用，并付诸实施。从而形成了美国文官考试录用制度。

法国的文官考试制度，是在英国效仿中国科举制建立了文官考试制度后，参照英国的文官考试制度建立起来的。早先是采用

间接考试制度，即考生通过考试进入有关学校，毕业后成为政府官员。在18世纪末和19世纪初拿破仑执政期间，规定录用官员与教育制度结合，毕业生先通过任职考试，再分别录用为各种不同级别的官员。由于学校教育的管理权集中在中央政府手中，教育程度与官员录用密切关联。因此，在学校录取考试与官员任职考试中，尤以第一种考试为重要。而进入高级文官阶层的人，必须进入国立行政学校学习。出于此种原因，入学考试竞争异常激烈，淘汰比例很大。但学生一旦被录取，实质上已进入法国文官系统。毕业生分配时，各部门一律录用。

从上述几国近现代文官考试制度的形成与发展中可看出，西方的考试制度深受中国科举考试的影响，正如孙中山先生在考察欧美诸国考试制度后所指出："现在各国的考试制度，差不多都是学英国的。穷流溯源，英国的考试制度，原来还是从我们中国学过去的"。①但上述各国的考试制度又各有特色，均与各自的国情一致。另一方面，各国文官考试制度的建立，又促使学校教育采用考试检查学生成绩和选拔新生，企事业单位通过考试招聘职员和选录工人。于是，考试作为甄别、选拔人才的手段被社会广泛运用，进而加快了考试科学的发展，考试因此出现了新的飞跃，形成了现代考试的模式。现以学校考试为例，叙述其演变过程。

在近代考试发展史上，17世纪世界著名的捷克教育家夸美纽斯对学校教育考试做出了卓越的贡献，他不仅建立普通教育考试体系，而且对大学的考试制度也提出了改革意见。当时，大学以不公平的考试为手段，排斥平民入学，以独断考试的行为授予学位。夸美纽斯则主张所有的人都应接受教育，学习"百科全书式"的知识，而且要学得"完善"、"彻底"，同时，他还提出

①转引自王康等：《人才知识手册》，湖北科学技术出版社1985年版，第69页。

了改革大学考试的主张：第一，举行公开的大学入学招生考试，择优录取勤劳、德行优良的学生，不允许那些靠世袭地位，只求安逸和奢侈不实的学生进入大学。要求考试公平，在分数面前人人平等的前提下，以考试成绩的优劣为标准录取学生。第二，学校考试不应该由某一个人独断，而应由学位委员会全体一致决断，把学位授予那些值得给予的学生。国家或国王应委派人员每年去大学检查一次。第三，学位考试采取答辩的方式。夸美纽斯的考试观点和改革考试的主张，实际是代表资产阶级及广大市民的要求，即变分化考试为统一考试，改变封建等级制度下剥夺人们应考权利和受教育权利的现象。这在当时的历史背景下，确有其进步意义，但也有其局限性。在广大劳动者没有解决温饱、人身权利毫无保障的情况下，劳苦大众要接受普通教育和高等教育，乃是可望而不可即的事。因此，它只利于资产阶级。但就考试发展而言，奴隶社会、封建社会的考试束缚、殆误人才，而统一考试有利人才脱颖而出，有利人的才智发展，并为考试自身的发展提供广阔的前景，也促进了社会的前进。从这个角度讲，夸美纽斯的考试思想奠定了近代西方学校考试的基础。在夸美纽斯考试思想的影响下，英国在1862～1870年间，私立学校每年举行一次统一考试，由政府派出观察员主持，根据考试结果分配补助金。美国也是如此。他们通过考试督促学校教育工作，评价教育质量，并根据考试结果确定财政拨款。这样，从19世纪起逐步建立了学校考试制度。这一时期，西方资本主义社会学校考试的特点是：

第一，各国政府通过考试控制教育，督导学校按资产阶级要求培育人才。当时学校考试有两种，即校外考试与校内考试。校外考试由国家教育行政机关或社会考试机构主持，校内考试由学校行政及教师掌握。英、法等国对校外考试特别重视。法国教育的管理权集中在中央政府手中，对学校课程、教科书等，均有统

一的规定，并通过考试检查学校执行政府规定的情况，不仅证书考试由政府视察员监督，而且连学生的平时考查、课堂作业也要经视察员检查。英国的国民教育体系是建立在地方分权的原则上的，国家教育厅没有规定强制的授课计划，只是规定各科最低限度的知识标准。因此，英国在19世纪末是通过各种考试来统一中小学教育标准，并根据在校学生数和考试成绩发给学校补助金。中学的校外考试在1858～1917年间，由社会考试机构主持，以后校外考试由政府主持。西方国家学校教育不管是由中央集权管理，还是由地方分权管理，它们共同的特点，都是通过考试检查学校教育质量或统一教育标准。由于校外考试严格，因此校内考试相应加强，教学也跟着考试打转，考试成了指挥棒，给现代考试留下不良的影响。

第二，校外考试与升学、就业、免费、奖学金等密切相联。因此，考试竞争激烈。学生要进入教育质量高的学校或名牌大学，必须经过激烈的入学竞争考试，而一般高等学校，在法国有“学士证”，在英国有“普通教育证书”或“高等普通教育证书”就可入学。但是要取得这些证书，同样需通过竞争考试获得。另外，学生在中学为了争取免费学习、在大学为了获得奖学金等，也必须参加竞争性考试或“奖学金级”考试。总之，19世纪学校考试的一个突出特点是考试竞争激烈。

第三，考试内容限于书本知识，重在死记硬背。如当时法国的“初等教育证书”考试，采用笔试和口试施测。笔试包括作文、默写和文法，数学计算题和度量衡制，史地常识等。口试包括阅读、文法、会话、背诵诗歌及口算等。试题由主管教育行政部门负责拟定，对学生保密，也对学校和教师保密。

第四，考试方法由中世纪的口试发展为笔试，而且以笔试为主。时至19世纪，随着班级授课制的推行和中小学学生人数的增加，笔试开始普及，到了19世纪末，笔试成为主要形式，口试居

于次要地位。

考试作为学校教学活动的重要组成部分，对教学质量的检测和督导教育目标的实现，起到了积极的促进作用。但是，考试在当时劳动人民生活毫无保障的情况下，对其子女受教育的权利却是一种排斥，对社会产生了一定的消极作用。

19世纪末20世纪初，美国兴起教育改革运动，亦称进步教育或新教育运动，其主旨是反对当时美国沿袭形式主义占统治地位的传统学校教育。由于这一运动的影响，当时美国形成了三种教育主张：（1）使教育成为一门象物理学或化学那样的科学，把学校变成一种工艺机构；（2）使每个儿童都充分自由地进行探索、发明、调查和创造的教育规划，每个儿童都是生活在实验室里具有创造力的研究员，形成以儿童为中心的学校。（3）教育要为建立一个合作的、有计划的社会创造条件，在这种社会中，人们将得到足够的社会和经济的保障，从而能自由地过着创造性的生活。即改造社会的学校。这三种主张的教育改革运动，对美国的学校发生了重大的影响，并波及欧美及亚洲各国，成为风行一时的新的教育思潮。它的特点是反对传统的学校教育，改革学制、课程、教学方法和教学组织，同时也对传统的考试制度进行改革。改革考试的主张是：

第一，认为考试主要不是同别人比分数，而是看学生的进步程度，可使学生看到自己的进步或不足，以激励他们作自我努力或自我竞争。因此，要重视校内考试，尤其是经常性考查。

第二，主张学校教育实行单轨制，放宽招生考试，使同一年龄的学生进入同一类学校。反对传统的竞争考试，使学生接受教育机会均等。

第三，主张通过智力测验，根据学生智力水平的差异分班分组，以根据学生的个别差异进行教育。反对传统考试仅考查学生的知识水平和机械记忆能力，并以考试的分数决定学生的升留级

的做法。认为这种考试没有区分学生的智力水平，而是要求学生齐步前进，故不适应不同学生的兴趣和能力，是违反学生身心发展规律的。

第四，认为传统的论文式考试存在严重缺陷。首先，对所考学科覆盖面小，测试内容局限性大，不能全面反映学生的知识和能力的水平。其次，评分不客观，不能准确评价学生的水平。因此，主张考试应推行"标准化测验"和非标准化客观测验。此种观点在美国影响很大，西欧和亚洲一些国家也深受影响。智力测验、标准化测验和非标准化客观测验从此逐渐兴起。第一次世界大战后，新教育运动由于被认为不能提高教育质量而逐渐衰落。智力测验和标准化测验由于滥用和自身的理论不完善及技术问题而遭到社会批评，而进入自我反省与完善阶段，即由科学工作者潜心研究和再行实验，以求理论完善和技术更新。后来经过不断改造，才形成了现行标准化考试的格局。现阶段，考试已渗透于社会各个领域。现行考试的类型五花八门，方法多种多样，大体上可归纳为三个类型：

（一）标准化考试

标准化考试在30年代受挫后，经过40年代反思、重新研究与实验，使其在理论上进一步完善，在程序上更为系统、科学，在手段上引进了现代科学技术的新成果，完成了手段上的技术革命。这样，标准化考试实现了高度数量化，在科学性、客观性和有效性方面，达到了前所未有的水平。同时，由于借助电脑技术，故大大降低了工作量，提高了考试的经济效益。标准化考试从此走上了健康发展的轨道，并为社会广泛应用。现今的美国，无论是社会系统的考试，还是学校系统的考试，基本上都是使用标准化考试。目前欧洲、亚洲、非洲一些国家和地区的社会和学校考试，正由分散到集中，积极进行考试改革，竭力追求考试标准化。苏联和东欧社会主义国家，也在吸收标准化考试的经

验，借以完善和更新现行的考试。

标准化考试是属于客观性考试的范畴，它可分为常模参考性考试与目标参考性考试。前者以常模为参照点，反映应试者成绩在考试总成绩中的位置，一般用于选拔、水平、安置等方面的测试。后者以专业或学科的目标为参照点，通常用于专业深度测试、学业成绩测试、合格考试等。

标准化考试是现代社会比较进步的一种考试，被视为考试改革的方向，但仍然有它明显的不足之处。为其模式所决定，标准化考试在发散性思维、创造性能力及表达能力的测试方面，远不及论文式考试有效。这些问题已引起科学研究者的关注，正在探索研究之中。

（二）论文式考试

论文式考试起源于古代文明社会，后为现代社会继承和发展。它的特点是：自由应答，有利于发散性思惟和创造性能力及表达能力的测试，但题量少，覆盖面窄，评卷标准难于掌握，误差不易控制。因此，考试的客观性和有效性程度低于标准化考试。加之这种考试的试卷目前尚不能使用光学扫描器和电子计算机评阅，投入的人力多，经济消耗也大。为使论文式考试扬长避短、适应现代社会发展的要求，人们正在进行积极慎重的改革。

（三）综合性考试

鉴于标准化考试与论文式考试各有利弊，人们便综合标准化考试与论文式考试的优点，取二者之长形成综合性考试。

程序、技术手段及组织管理实现现代化，考试试题

的，采用客观试题与论文式试题相匹配组

准化考试和论文式考试不足之处。有

局。第一次考试为合格考试，测

规定标准，如日本高考的学

《高中学习指导要领》

次考试，是主试单位根据其需要，测试应试者的专业适应能力。

总之，考试的统一与飞跃是现代文明社会考试的特点，也是考试发展的必经历程。考试的统一是考试面向全社会，破除古代社会考试的等级限制。考试的飞跃是考试内容、方法、技术手段的现代化，而考试的本质仍然是甄别、测度人的德、学、才、识、体之个别差异。

综上所述，我们认为，考试是在人类历史的长河中发展起来的，它受制于社会的政治经济，又有其自身发展的规律，即必须根据社会的要求和人的年龄特征，选择考试的内容和方法，运用科学的技术手段，才能取得最佳的测试效果。同时，任何一个社会的考试，必然是继承前一社会考试的精华，摒弃糟粕，运用当代社会的科研成果（包括新的考试理论和先进的测试手段及其他技术设备）形成适合当代社会发展需要的考试模式，才能发挥时代所期求的功能，这就是考试发展的必由之路。

四、考试的变化趋向与模式

现实世界是两种社会（社会主义社会和资本主义社会）制度并存．无论哪种社会制度的国家，都面临着新技术革命的挑战。关于人类社会发展的趋势，从社会制度来说，将从资本主义经过社会主义达到共产主义。从社会生产和生活来讲，将实现生产和生活的全面自动控制，把人类生存的领域扩大到地球深处和外层

人类对物质生活和精神生活的要求将越来越高，人类也将

自觉地更有计划地管理自己。而当今世界生产发展的形

[illegible]；生产组织既严密又灵活；电子计算机综

[illegible]；信息已成为整个组织的中心环节；

[illegible]要的学问；体力劳动者比例下降，

[illegible]重要性更突出。根据有关研究

[illegible]代，社会所需求的人才，

将是智能型、创造型、社会型“三型合一”的人才。教育将是通才教育和英才教育，考试将由现代固定型的静止考试，发展为变通式考试，由整齐划一的团体考试发展为因才施测的个体考试。由主试和被试对峙的考试发展为现代人才的能力模型为主试、由本人进行自我测度的考试。考试手段将向全自控化、信息化发展，应用智能化的机器人进行人机对话将是未来考试的基本方式。研究未来、预测未来是科学发展的动力。对考试未来发展趋势与模式的预测，也将鞭策人们深入开展考试科学研究，以适应考试科学的发展，使考试理论、方法与技术的发展，跟上时代发展的步伐，充分发挥考试对人类社会发展的积极作用。

第三章

考试功能

> “活的机体，只要具有多种功能，并且善于适应环境的变化，就会蓬勃发展起来。考试之所以蓬勃发展，同样是由于它具有多种功用的性质。”
>
> 〔英〕罗伯特·蒙哥马利

相对人类社会的整个历史进程来说，考试的使命还远未完结，它在人类社会进化中充分施展其固有功能还仅仅是发端。

长期以来，对于考试的功过，历来乃仁者见仁，智者见智，素无定评，颂扬者颇多，批驳者非少，主张废弃者历代亦有之。然而，在科学技术发展速度空前、人类生活和生产方式急剧变革的今天，“考试已经这样稳固地站定了脚跟，要废除它似乎比取消篝火节或者圣诞节更无可能”①。它不仅被社会生产和社会生活的各个领域广泛采用，而且成为当今世界人们热切关注、苦心探讨的一门新兴科学。为使考试适应现代社会日益变化的需求、迎接富有挑战性的未来，许多国家不惜人力、物力、财力的投资，去寻求和揭示考试的规律，对考试的模式不断地革旧更新。

基于上因，生活在20世纪末这个不寻常时期的考试研究工

①〔英〕罗伯特·蒙哥马利著，黄鸣译：《考试的新探索》，广西人民出版社1984年版，第76页。

作者，无论从何种角度研究考试与人类物质和精神文明的关系，都将会面临并必须回答这样一个问题：考试在人类社会发展中的功能何在？

对于考试的功能问题，古今中外不少学科的专家、学者均有阐述，只是由于受限于各自研究的对象，不可能得出一致结论；又因考试自身发展和人们对考试现象认识上的局限，从而导致对考试功能认识的偏颇，认为考试的功能只是体现在学校教育系统，其它社会系统即或有所反映，也不过是用于人才的选拔而已，故凡论及考试功能大都是学校考试的功能。此类观点在国内外有关著作和文章中多有所见。日本教育界著名学者梶田睿一在《现代教育评价论》一书中写道："无论是考试还是考查，都是为了把教育本身放在科学的合理的基础上，从而，就全局来说，运用更合适的形式促进儿童、学生的成长、发展，"又说，以考试为中心的评价形式，"无论是把它看作反馈的机能也好，或是看作整速器的机能也好，都是学习者为了以稳健的步伐、有组织地进行学习所必不可少的措施。"香港学者肖炳基在《教育测量与标准化考试》讲义中指出："考试是了解学生的程度，发现学生的优缺点，以审查或预估学生的发展。"苏联的舒金娜从学校教育的四方面论述了考试的功能，即学习功能、组织功能、发展功能、控制功能。游正伦在《教学论》一书中说，考试"对学生来说，鼓舞学习，督促复习；对教师来说，检查教学，了解效果，改进教学，总结经验；对学校及上级教育机关来说，了解教师教学情况，教学工作问题，改进领导；对国家来说，是选拔人才，择优的手段。"以上不同国度和时期的考试研究者对考试功能的阐述，仅就学校教育系统考试的功能而言，只阐明了考试在学校教育这一极为有限范围内的功能，而忽略了考试在整个人类社会中的作用。事实上，考试的功能所及，远不是考试工作本身，也不止是主试、被试及其家庭的范围。

随着考试自身的演变发展，以及人类社会使用考试范围的不断扩大，“考试的功用甚为广博。……在政界、军界、工界或商界各方面，应用考试来选择人才的，一天普遍一天，而且甚富成效。可见考试不特为解决各种教育问题的必要工具，且是一种推进各种社会事业的良好法则”①。英国当代考试研究专家罗伯特·蒙哥马利在论证考试功能时指出：“活的机体，只要具有多种功能，并且善于适应环境的变化，就会蓬勃发展起来。考试之所以蓬勃发展，同样是由于它具有多种功用的性质。”②他认为考试能适应人类社会多种目的需要，其功能也是多元性的，能对社会结构的不同方面产生影响。他在所著《考试的新探索》一书的第一句话就明确写道：“考试对于决定英国社会的结构，起到了微小却又极为重要的作用，如同发展中的社会舞台一定显示出的简单轮廓那样。”我们认为，凡存在刻度、判别人的德、学、才、识、体等个别差异客观需求的地方，就必有测度、甄别人的个别差异的考试现象，正如哈佛大学校长罗惠尔在论及考试的功用时所云：“每种职业都含考试在内”③，而“每种考试几乎都包含一种以上的功用”④。考试作为人类社会母系统中一个相对独立的子系统，有不同于其它子系统的多种固有功能。在人类社会中，考试是实现教育目标、开发和合理利用社会人力资源的有效措施，是促进社会发展的重要机制，其主要功能大致可归纳为三大方面：调节功能、行政功能和督导功能。

①〔美〕亚尔保德·兰著，浦漪人、黄明宗译：《新法考试》，正中书局1935年版，第5页。

②〔英〕罗伯特·蒙哥马利著，黄鸣译：《考试的新探索》，广西人民出版社1984年版，第14页。

③④〔美〕亚尔保德·兰著，浦漪人，黄明宗译：《新法考试》，正中书局1935年版，第30、16页。

第一节 调节功能

如果说，"凡是具有交互与共同关系，与表现交互与共同行为的一群人，都可成为社会。"③那么，人类社会便是一个多元、多维、多层、多要素交互综错的结构网络（或组织系统）。这个巨大的母系统包含着若干分系统，每个分系统又可分为众多的子系统，如图3－1所示。无论是分系统与分系统之间，还是分系统与子系统之间，以及各子系统之间，都存在着相对稳定的内在关联。它们互为条件、相互调节、彼此控制，各自以特殊的结构形式发挥其在母系统中的独特功能，维持整个母系统的正常运行，推动社会协调发展。

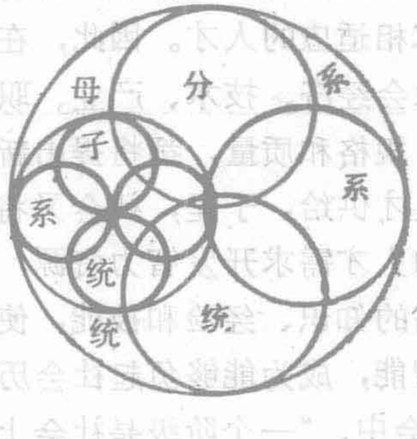


图3－1 社会结构系统简图

在人类社会母系统中，人才群体是一个极为重要的分系统。它的结构及其功能，对整个社会系统具有牵制作用。人才群体系统，是育才（教）、量才（考）、用才（用）三大子系统的有机组合，其合理结构的形式和功能的有效发挥，必以三个子系统的

③复旦大学分校社会学系编：《社会学文选》，浙江出版社1981年版，第149页。

功能和条件“耦合”为前提。在人才群体系统中，教、考、用三个子系统的条件和功能是相互依赖、互为因果的，其中任何一个子系统失去效用，都会破坏另外两个子系统的条件和各自功能的发挥，进而影响人才群体系统的整体效能。考试在人才群体系统中处于中介地位，对育才和用才两个子系统具有调节和控制的杠杆功能。这突出反映在两个方面。

一、人才供求标准的调节器

人是生产力的本质要素，是推动社会发展的基本动力。在人类社会的一切财富中，尤以人才最为珍贵。无论是在奴隶社会、封建社会、资本主义社会，还是在社会主义社会及其高级阶段的共产主义社会，要想不断丰富社会物质财富和精神财富，以满足人类日益增长的物质和文化生活的需求，都有赖于足够数量并与其社会物质生产要求相适应的人才。因此，在人类社会发展的不同历史阶段，随着社会经济、技术、产业、职业结构的变更，对人才的数量、种类、规格和质量，都将提出新的需求。有新的人才需求就要有新的人才供给，于是，社会又将赋予教育新的历史使命：根据社会新的人才需求开发智力资源，授予新生一代社会生产与社会生活必需的知识、经验和技能，使之具备与其社会要求相适应的素质和智能，成为能够负起社会历史重任的新型劳动者。然而，在阶级社会中，“一个阶级是社会上占统治地位的物质力量，同时也是社会上占统治地位的精神力量。支配着物质生产资料的阶级，同时也支配着精神生产的资料。”①教育也是如此，其支配权同样为支配着物质生产资料的阶级所控制。当社会生产力的发展提出新的人才需求之后，并不等于教育就一定能按需培养出合格的人才。

①《马克思恩格斯选集》第1卷，人民出版社1972年版，第52页。

一方面，人类社会一定历史发展阶段提出的新的人才需求，必须被从事社会实践的人们感受和认识，在此基础上形成新的人才观，进而促使教育观、教学观的更新，并为一定社会支配物质和精神生产资料的阶级所接受，由他们根据社会人才需求、本阶级利益及社会能够提供的育才条件，确立具体的育才对象、内容、目标，制定相应的教育制度、方针、政策，以及行之有效的考选措施，才能作用于人才培养，进而实现因需设教、按需育人的目的。

另一方面，人才培养具有"持久性"的特点，"十年树木，百年树人"，"因要改变一般的人的本性，使它获得一定劳动部门的技能和技巧，成为发达的和专门的劳动力"①，通常要经过十几、二十年的学校教育。其间，学校教育培养人才的标准处于相对稳定状态，而社会对人才素质与智能的要求则随生产力的发展和科学技术的进步不断提高，不会始终停留在同一水平，这就难免要产生人才供求标准不相一致的矛盾。因此，在人才培养的过程中，为使所育人才的规格、质量切合社会需要，除创设科学的育才条件外，还必须有既反映社会人才需求又能检验人才培养标准的科学手段，根据社会人才需求变化不断调节人才培养标准，借以维持人才供求标准动态的平衡。

诚然，"实践是检验真理的唯一标准"，也是检测人才质量的客观标尺。学校培养的人才究竟符合社会需求与否，最终要在社会实践中进行鉴别。但这种检测手段，只有在新生一代已被培养成"准人才"②且输送到社会劳动部门后才能奏效，所以不能及时作用于人才培养，敦促各级教育部门随即调整人才培养标准，提高教育的内部效率，增强人才的社会适应性。因此，当人才尚未踏入社会之前，即在培养阶段，人们采用了另一种贯串人

①《马克思恩格斯全集》第23卷，人民出版社1972年版，第195页。

②准人才——"指基本上具备人才要素(德、学、才、识、体)，能基本适应某项工作者．如大学毕业生"便属此类人才(《人才知识手册》)。

才培养始终的检测手段——考试，由它充当社会人才供求标准的调节器。

考试在社会人才供求标准之间的调节作用，主要体现在两方面：

其一，促进教育机关根据社会提出的人才标准培育新生一代。

先说人才及其标准问题。何谓人才？现阶段并无公认的界说。我们认为，所谓人才，简而言之，就是人群中较之同行更富智慧与才华的人。他们总是以某种方式存在于一定的社会中，是一定社会历史阶段的产物。如果"人才离开了他的社会有机体，也就不成其为人才了。"①所谓人才标准，亦即一定社会对所需人才在德、学、才、识、体诸方面的具体要求。

由于一定社会的人才标准，是为一定的生产方式所制约的，不同历史时期的社会经济、文化、科技发展水平和统治阶级对各类人才德、学、才、识、体各方面的标准不尽一致，德、学、才、识、体的具体内容及社会对人才德、学、才、识、体要求的侧重面也各不相同。因而，不同社会及其历史时期的人才标准，无不打上所处时代的烙印，具有特定的本质内涵。

但是，不同社会或历史时期人才标准的特殊性，并不排除人才标准组成要素和实现人才标准的某些共性特征。比如，德、学、才、识、体几方面的要求，是组成一定社会人才标准的基本要素；教育和考试，是实现和调节人才标准的基本手段，等等，都未因生产方式的变更或时代的发展而有别。

考试之所以能够促进教育根据社会提出的人才标准培育新生一代，原因在于考试与教育既有目标上的一致性，又有职能上的各自独立性。

考试和教育因人类社会发展的人才需求而产生，又以满足社

①《智力开发》，上海交通大学出版社1985年版，第63页。

会发展的人才需求为目的。一定社会的教育目标和考试目标，是社会对人才要求的具体反映。培养、选拔什么样的人才，其标准均以一定社会对人才德、学、才、识、体等方面的要求为依据，这是教育和考试的普遍规律。在人才培养的过程中，考试与教育在目标上的一致性，是通过促进教育按照社会提出的人才标准育人，以及检验培养的人才是否达到社会提出的人才标准而实现的。无论是诊断考试、合格考试，还是选拔考试，其内容和标准，都反映了某种课程、教学和教育目标的要求，体现了某种课程、教学和教育的基本任务和目标。各种考试的结果，它不单是显现学校教育培养人才的达标度，而且暴露出教学中存在的问题和所育人才与社会需求标准之间的差距，构成调整、改进教学的依据之一，故考试可作用于人才培养的标准，促使人才培养趋向社会需求标准。同时，凡符合社会、教育和人的成长规律的考试，可通过科学地测度和甄选为人才创设达标、争优的条件与机会，使人才的成长始终处在提高素质与发展智能的流动性竞争状态，优秀者被竞争步步推上高一层次的人才培养阶段，而那些未达标的人，又被竞争推回到其原居层次重新接受教育。这样，人才培养过程变成了人才流动过程，随着考试的调节，人才在连续不断的竞争中得到提高，促使他们向社会提出的人才标准进取。

再者，由于学校考试制度既是教育制度的重要组成部分，又是国家考选制度的重要组成部分，学校系统举行的各种考试，既要服从学校教育的基本任务和目标，保持目标上的一致，同时又必须反映变化发展的社会人才需求，促进教育水准不断上升。因此，考试与教育的职能又是相互独立的，教育旨在根据社会需求育才，考试则根据社会需求检才和选才。而任何一种考试又无不反映某种社会的人才需求，其测试结果总是含有社会评价性质，人才符合某种考试标准，亦即社会对教育成果及人才某方面素质与智能的承认，能给教学双方以强烈影响。正由于此，考试

可通过反映社会变化发展的人才标准，调节人才培养的标准，以维持人才供求标准的平衡。

其二，调解人才供求标准矛盾，促使需要与可能有机结合。

社会人才供求标准的变化发展，如同世间其它事物的发展一样，平衡是相对的，不平衡则是绝对的，人才供求标准之间的矛盾客观存在。在现代社会中，人们为使人才培养在数量、种类、规格、质量方面适合社会需要，大都采用人才预测的办法解决人才供求矛盾。但人才需求总是先于人才供给，人才培养必然后于人才预测，求与供之间，不仅会因人才培养的"滞后性"产生标准上的差异，而且会因人才培养的条件限制，造成育才可达标准与社会发展需求标准的差异。诸如，当人类社会跨入工业社会后，新的生产力要求有数以百万计的、既懂机器原理又能正确使用机器的生产工作者，由其承担现代生产的使命。为满足这一社会需求，当时的学校教育随之改变了原来的人才培养方式，将师传徒代之以"批量生产"，但并未因为人才培养方式的改变就能立即从数量和质量上提供完全符合社会需求的人才，而是在一个相当长的时期内才逐步适应的。又如，随着近、现代科学技术的飞速发展，现代生产进入了新的发展时期，智力和知识成为现代生产者劳动技能的基本要素。它要求现代人才具有正确驾驭和灵活运用所学知识的能力，富有开拓与创造精神，从而迫使教育摆脱"仓库理论"和"填鸭式"教学方法的束缚，将开发智力、培养能力置于人才培养的首要地位，为社会提供开拓、创造型的现代人才。但这一现代人才的标准，不是学校教育一下子就能全部达到的。对于社会的人才需求，作为教育理应有求必有供，但并非有求就能供。一种新的人才观的产生，人们需有一个认识和适应的过程，学校培养和造就一代社会发展所需的新型人才，在育才条件方面，也需有一个相应准备、改进或创设过程。这就难免出现人才供求标准不平衡的矛盾，要求人们在人才供求过程中采取有

效的调节措施，使需要与可能有机结合。

在中外历史上，解决人才供求矛盾的方法有不同之处，更有共同之点。所不同的是，因社会制度不同而采取的方法不一样。在资本主义国家里，是通过劳动力市场的自发调节来解决的。在社会主义制度下，则是根据社会主义有计划按比例与适应社会发展的人才需求相结合的原则解决的。但是，对于人才供求矛盾中的人才标准的调节，不同时代的人们又是通过同样的途径来实现需要与可能的结合的。考试，就是人们常用的基本措施。

在社会人才群体系统中，考试是供求双方的中介，其标准既反映社会人才的需求，又要顾及人才培养的可能。当社会对人才标准的要求超越人才培养实际能够达到的水平时，考试便适当降低社会需求标准，以人才培养已经达到的最高水准为据，确立人才评量或选拔的标准；若人才预测科学而准确，育才条件优越，使人才培养有一定的"超前性"，人才具备的实际水平高于某些产业或行业的人才需求标准，考试又将以社会需求的最高标准为据，确立人才定级、择优的标准。如此循环反复，不断调节供求双方的矛盾，使需要与可能有机地统一起来，以维持社会人才供求标准的大体一致。

二、调节社会各部门人才结构的杠杆

考试调节社会各部门人才结构的杠杆功能，突出反映在两个方面：一是协调各部门的人才结构；二是影响各部门合理人才结构的形成。

这里所说的社会各部门，系指从事物质生产、直接创造物质财富的一切企业和单位，以及服务于社会生产与社会生活、不能直接创造物质财富的一切企业和单位，亦即社会生产部门和非生产部门下属的所有部门，诸如工业部门、农业部门、林业部门、渔业部门、畜牧业部门、商业部门、文教部门、科技部门、卫生

部门，等等。各个生产部门和非生产部门，又因劳动过程的社会结合、活动的内容及内部分工的不同，而分化为若干类别和多种层次的部门或单位。这些部门或单位，在人类社会“生产力不断发展，社会分工愈来愈细、生产日益专门化”的过程中逐步形成，并随人类社会科学技术的进步、生产力水平的提高、社会分工的变革，而不断地发展和再分化。

社会不同生产部门与非生产部门的形成、发展和不断分化，既是社会生产力日益发展、生产专业化程度不断提高、社会分工逐步加细的必然结果，又是人们征服自然、发展生产力、推动社会前进的必然要求。人类社会的生产和生活自始至终都是社会性的，每一社会成员都以一定的劳动方式作用社会和影响他人，同时依赖社会其它成员的劳动成果维系自身的生存和事业的发展，恰如《管子》所云：“一农不耕，民有为之饥者，一女不织，民有为之寒者。”或谓“盖一人之力，可以影响多人之生；反之，一人之需，有恃于多人之力。”①此种人与人之间的彼此依赖关系和交互作用，便是社会生活与生产赖以形成、发展的基础。

但是，人类要征服自然，推动生产力发展，不断丰富物质财富和精神财富，使社会日趋发达，满足人们变化发展的物质、文化生活需求，仅有每个社会成员的孤身劳作及个体间的交互依赖关系是不成的。《荀子》曾云：“人，力不若牛，走不若马，而牛马为用何也？曰，人能群，彼不能群也。”对此，恩格斯也曾作过科学论证，他指出：“许多人协作，许多力量溶合为一个总的力量，用马克思的话来说，就造成‘新的力量’，这种力量和它的一个个力量的总和有本质的差别。”②单个的社会成员，只有根据社会

①复旦大学分校社会学系编：《社会学文选》，浙江人民出版社1981年版，第141页。

②《马克思恩格斯选集》第3卷，人民出版社1972年版，第166页。

经济发展和社会分工的需要，组合成具有一定组织形式的群体，并仰赖与社会其他群体的共同关系和交互，结成整个社会成员的群体网络，才能使社会各成员的智慧充分表现，并维系他们的交互与共同关系，进而构成征服、改造自然的巨大社会功能，实现改变现状、创造未来的目的。

可见，每个社会特定范围内的单个社会成员都不是孤立生存的，总要与同一社会的其他成员发生交互和共同关系，并组合成各种不同类型、不同层次、具有一定组织形式的群体，亦即一定社会中内容、性质、规模有别的部门或单位，如现实社会中的工厂、学校、军队、商店，以及工厂中的车间、班、组，军队中的师、团、营、连，行政机关中的处、科、室等，都是由单个社会成员组合而成的不同类型与层次的群体。同时，工厂之所以成为工厂，学校之所以成为学校，不单因为工厂、学校是由不同的社会成员组合而成，还由于工厂、学校内部的各成员之间有一种特殊的结构，供其内部各成员表现与维系他们的交互和共同关系。

作为较之一般社会成员更富有才能的人才，同样不是"一盘散沙"。他们也是以一定的方式存在于社会的各个部门中，总是根据不同部门或单位的客观需求，按照一定的组合方式构成相应的人才结构，以维系同一部门人才之间的交互与共同关系，让每个人才表现其才能。任何一个部门，其工作效率和社会效应都取决于这种人才结构的合理与否。就是说"量的自然的、机械的积累，并不必然引起事物的某种质变。……只有当他们有机地结合在一起时，才会产生力量，才算是真正的生产力。"①结构排列组合不同，产生的功能就不同，其效应也不一样，这是自然、社会和精神领域普遍存在的现象，如同样的机器构件不同的组合方式，可产生不同的功能；同一金属体内分子的排列组合不同，其

①光明日报社主办：《文摘报》，1987年第415期第6版。

强度就会出现显著的差异；一个部门的人才结构与功能也是如此，马克思在《相对剩余价值》一文中举例说：12个人在144小时共同工作中提供的总产品，比12个人单干的劳动日每人劳动12个小时或一个人连续劳动12天可提供的产品多。这说明同等数量、素质与智能的人员，由于组合方式的改变，所产生的功能及其工效是不同的。因此，人们在创造物质财富和精神财富的过程中，为求用同等的人力、物力取得最佳效率与效益，总要依据生产或工作需求的变化，经常调节部门内部的人才结构，使之保持合理，趋于优化。

社会各部门人才结构合理化或优化问题，实质上是部门人才结构诸要素组合的科学化问题，即人才在年龄、专业、知识、智能、体能、性格等方面的比例、层次是否协调，是否符合部门的客观实际。当然，每一个具体的企业、部门或单位，在人才年龄、专业、知识、智能、体能、性格等人才结构要素的比例、层次和组合方式方面，都是各具特色的。但是，由于社会是以共同的物质生产活动为基础而相互联系的人们的整体，在社会这个巨大的母系统中，任何一个部门或单位的存在和发展，都必须依靠与其他部门的交互和共同关系。因此，不同部门或单位在合理人才结构的构成方式方面，又有许多共同的特征。一般说来，一个人才结构合理的部门或单位，它都是一个多序列、多层次、多要素的人才动态综合体，其人才的组合具有如下相同特点：所需专业人才种类齐全，综合交叉，比例适当，成龙配套；人才年龄结构完整，老中青人才有机结合，承上有师，启下有人，后继有梯队，无断裂层，有利于协同作战；人才的知识水平呈宝塔形立体结构，纵向高、中、低层次分明，比例合理，横向特色各具，组合严密；人才智能多型化，各类兼备，配置合理，“专能”可扬，“多能”可聚；人才性格兼容并包，彼此互补，协调一致；人才体能各殊，从业人才的体型、体力、适应能力等，与部门或单位

不同职类、工种、工序及环境相适应，在整体上是一个运动和谐的机体。

不同部门或单位的合理人才结构，不同组合方式具有同性功能，如现代国内外学者所公认的合理人才群体结构的"感应功能"、"互补功能"、"师承功能"、"协调功能"，等等。也就是说，一个部门、一个单位效能的优化，取决于部门或单位内部人才结构的优化，而社会各部门人才结构的优化，又取决于部门人才结构组合与调节手段是否科学。

在人类社会的发展过程中，一切社会分工和职业结构，总是随生产力水平的变化而变化，为社会分工和职业结构所决定的部门人才结构，同样是一个动态的综合体。一方面，原有人才自然减员（病故或自然死亡），新生人才要相继充实，这种人才新陈代谢的自然现象，将不断改变着各部门的人才结构。另一方面，随着部门事业的发展，分工的变化，或因技术更新、人才流动等原因，又会导致另一种人才陈新代谢的现象，即淘汰不合格人才或多余人才，引进合格的新人才，这也将改变各部门的人才结构。所以，社会各部门的人才结构总是处在变化、发展的运动过程中。为保持部门内部人才结构动态的平衡，尚需不断地进行调节。然而，任何有效的调节都必以科学手段为前提，只有在明了本部门人才客观需求的基础上，切实解决和正确使用选贤任能的方法，才能真正达到预期的调节效果，实现按需择人，因才施事，人与人相容，人与事相配的目的。

当然，人才具有多样性，俗话说："三百六十行，行行出状元"，不同类型的人才又有不同的特点。因此，调节部门人才结构的方法应是多种多样，理当不拘一格。比如，贤能荐举，群众推荐，民意择选，个人自荐，领导察选，考试选拔等等，都是人们选任贤能、调节部门人才结构常用的方法。但尤以考试的使用最为普遍，它一直是社会各部门调节或形成合理人才结构的基本

手段，富有调节功能。

（一）考试对维持部门人才结构动态的平衡具有协调功能

调节一个部门的人才结构，旨在谋求人与人和人与事配备合理，使人才的年龄、专业、知识、智能、体能、性格等，与本部门工作职位的总体要求协调一致，有利于一个部门最佳整体功能的发挥。一个部门人才结构的调节过程，既是对人才"检测——选拔——任用——再检测——再选拔——再任用……"的反复循环过程，也是使人才结构变化、发展的运动过程，成为"由不合理——合理——再不合理——再合理……"的良性循环过程。

人们常说："善任必先知人"，"施事贵在因才"。这表明，合理使用人才的关键，就在于科学量才和选才。从调节部门人才结构的角度看，也可谓切中要害。在人才结构调节过程的三大环节中，检测是前提，选拔是基础，任用是目的，一般是前一环节制约后一环节。没有对人才客观准确的甄别和测试，就不可能对人才作出可靠有效的选拔，没有对人才准确可靠的甄选，也就难于达到恰当任使人才的目的。这就赋予考试以特殊使命，即根据本部门不同工作职位对人才德、学、才、识、体的要求，制定出各种测试标准、内容和方法，对在职人员的德、学、才、识、体等方面，分别进行定期不定期的考查、考核、或测试，客观掌握不同人才的实际水平，并依据考试结果进行人才结构调整，对不同素质与智能的人才作统筹有效之用。

考试是人类社会识才、辨才、量才、选才的基本方法，也是调节部门人才结构的主要手段，其协调功能客观存在，这是无可厚非的。但是，考试的质量有优劣，作用也有好坏，其协调功能的二重性同样不可否认。

长期的考试实践表明，凡标准客观、内容科学、方法得当的考试，不仅能如实反映一个部门对人才的需求，而且能比较客观

地检测人才在德、学、才、识、体等方面的个别差异。此种考试，对于一个部门平衡调整人才结构来说，它所产生的功能既是良性的，也是多方面的。其一，由于手段科学，就能较为准确地量才和选才；人才选得准，便可用得恰当，即做到按能就职，因才用人，使每一个人的工作职位符合自身的实际，其才能可得到充分的施展，进而收到人尽其才、才尽其用、事竟其功之效。其二，由于考试的标准和内容反映了部门对人才的客观需求，在调节的过程中可依据考试提供的信息按需择人，使人与事合理相配，实现人才结构在专业、年龄、知识、智能、体能、性格等方面的协调化，避免部门整体功能单一化的弊端。其三，由于考试标准如实体现了不同工作职位对人才的要求，它既是检验人才实际水平的尺度，同时又是不同人才工作进程上的奋斗目标，往往激励人才求知、进取，促使他们在同行间展开竞争，不断更新自己的知识、增长自己的才干、强化自己的适应能力，从而形成良好的“群体感应效应”，迸发出新的群体功能，推动部门事业的发展。

与此相反，凡是标准、内容、方法不科学的考试，它对调节部门人才结构的不良影响，也不是单方面的。如果考试的标准不符合工作职位的要求，测试内容与欲测内容相悖，测试方法与被测对象不适，通常会造成人才的三种浪费：人与事不配，能与职不称，虽占其位，却不能司其职、尽其责，虽耗财、物，而不能取得实效，这是第一种浪费；小才大用，力不胜任，妨碍事业进程，贻误时机，削弱整体功能，这是第二种浪费；大才小用，用不因才，尽管人才各有其位，也不免会出现有才难施或不能施的境况，造成人才过剩、怀才不遇的局面，其结果必然是“龙多不下雨”，这是第三种浪费。因此，在调节部门人才结构的过程中，要想提高人才使用的有效性，就务必注重量才、选才手段的科学性，切忌脱离实际、随意施考，甚至误用或滥用考试。对不

同类型的人才，以及同类人才不同方面所进行的考查、考核或测试，必须针对不同的考试对象和目的，精心设计出不同的标准、内容和方法，并严格按照不同考试的特点与程序实施。否则，就不能正确发挥考试的协调功能，达到预期的调节目的。

（二）考试对各部门合理人才结构的形成具有控制功能

常言道："流水不腐，户枢不蠹"，"生命在于运动"。一个部门的人才结构也是如此，它并非死水一潭，而是一个活的机体，并在人才新陈代谢的运动过程中经常获得新鲜血液的补充，不断增强自身的生命力。

众所周知，无论是生产部门，还是非生产部门，其规模、任务与目标，总是随社会发展而变化的。作为以适应本部门任务和目标需要而建立的人才结构模式，也必然随外部形势和本部门任务、目标的变化而变化。另一方面，随着时间的延伸，部门工作职位的变更，以及部门事业对从业人员知识水平和业务能力要求的提高，人才必然出现内外流动的现象，这同样会造成人才结构的改变。对于一个具体的部门来说，不管是上述哪种因素引起的人才结构变化，都有一个人才的新陈代谢问题，并面临着改造旧结构、形成新结构的任务。

在改造旧的人才结构的基础上，要形成新的人才结构，并使其处于最佳状态，首要的是物色、甄别、选拔人才应采取适当的方法和途径，检验他们实际具有的知识水平、业务能力、思想素质与体能，做到心中有数、因才用人。这一问题的解决，不仅要破除"任人唯亲"、"任人唯资"、"任人唯顺"、"任人唯全"等陈腐的辨才、选才观念，而且要有正确的人才甄别、选拔途径和有效的把关措施。考试就是人们常用的主要方法之一。

在部门人才结构的调整过程中，考试之所以被人们持久、广泛的使用，其重要原因，就在于它对部门合理人才结构的形成具有以下独特的控制功能：

1．考试对人才的检测是多层次多方位的，有利于选才质量的把关。

从方法的角度看，考试的形式是多种多样的、如笔试、口试、示范表演、实际操作、设计、答辩、接谈、竞技比赛、演讲，等等。这些方法既可分别采用，也可综合运用，对人才既可直接测试，也可间接测试，适应面十分宽广。从内容的角度看，考试的灵活性很强，通常是因人因事而异，对人才不同的素质、知识和不同层次的能力，可纵横兼顾，进行多层次、多方位的检测。从手段的角度看，考试能将定性、定量分析有机结合、在分析比较中互为补充，克服单一手段的不足，对人才作出比较客观中肯的评价，为准确选才提供可靠依据，进而也为部门合理人才结构的形成奠定了基础。

2．考试有一个相对统一的客观标准，有利于维护人才选拔的公平性。

考试对于人才的选拔，除在质量上把关外，还能在一定程度上抵制不正之风，清除“任人唯亲”等陈腐观念的干扰。正如人们所总结的那样，考试既是一把尺，又是一道门坎，差尺短寸者，翻不过门坎的人，都不在入选之列。因而，考试基本能保证所选之才就是欲选之才，也是能用之才。这对一个部门的人才调整来说，实际上起着控制人才质量的作用，使进入部门工作职位的人才，确系新鲜血液，并能与原有血液溶为一体，不致与部门事业需求失去平衡，从而形成良性循环，赋予部门人才结构以新的生命力。

第二节 行政功能

在现代社会中，考试特别是资格考试、选拔考试和研究性考试，已为社会生产和社会生活各种组织普遍采用。它作为一种培

养、检测、甄选、任使、管理人才的方法，人们已有深切的了解。对于考试在现实社会各项管理中的行政功能，也有不同程度的认识，如英国考试研究专家罗伯特·蒙哥马利就曾论述道："切不可以为大学生的取得资格考试的成绩是徒有其名的。在上一个世纪，各种行业吸收成员的测试早就进行了。对要求充当医生、药剂师、工程师或者其他专家的青年，都必须考查他们的才能。这样，当你需要他们提供服务的时候，知道这些专业人员的业务能力至少达到某种最起码的水平，你才会感到放心。业务能力达不到一定的水平，是不能充任教师和医务人员的，这好像驾驶执照的考试一样，就是要保证汽车司机具有起码的驾驶技术和知识。"①在这里，罗伯特·蒙哥马利已从一个侧面揭示了考试在现代管理中的行政功能，即考试是社会各行各业选拔从业人员的基本手段，为各类人才的恰当任使提供科学依据。但这远不是考试行政功能的全部。

实际上，考试在人类社会中的用途是极为广阔的。它是人类社会有效开发和利用人力资源的主要途径，是人类社会进化的重要机制，是社会一切管理必不可少的手段。因为人是社会的主体，人力是推动社会发展的基本动力。在人类创造物质财富和精神财富的过程中，如果离开人这个主体，没有人力的科学利用和正确发挥，一切资源都不过是客观存在而已，它不可能转化为现实的生产资料和生活资料，也不能形成改造主客观世界的活动，那么，社会发展必将终止。正是在这个意义上，国内外学者才明确提出了以人为管理核心的观点。他们认为，各个社会部门的管理工作，只有抓住人这个管理的核心，使社会的人力资源得到有效的开发和合理的利用，才能实现管理的目标，推动社会前进。既然人的管理是各项管理的核心，那么，人才的培养、甄选和任

①罗伯特·蒙哥马利著，黄鸿译：《考试的新探索》，广西人民出版社1984年版，第15页。

使，就自然是一切管理部门面临的共性问题，也就必然要借助考试这一手段，对各类人才进行检测、甄选、任使、调节和控制，以求人力资源的有效开发和合理利用，不断提高人才培养与使用的经济效益。因此，对于考试的行政功能，我们可从人事管理和教育行政管理两方面进行探究。

一、人事管理的重要手段

综观人类社会的人事管理，虽然名目繁多、内容有异、形式多样，并常随时代的变迁、统治集团利益与意志的不同而变化，但归根结底是对社会人力资源的开发和利用，它必以人才的培养、甄选、任使、考核、晋升、奖惩、调迁等为要务。而上述各项任务的完成，又必以了解和掌握人才的德、学、才、识、体诸方实际情况为先决条件。一个对人才全然无知的人事管理部门，不可能做到辨才客观、选才准确、用才恰当、晋升有据、奖惩有理。所以，自古以来，人们总是将"了解人"置于人事管理的首位，视为人事管理成败的关键。

所谓"了解人"，即是对人才的考试、考核、考察或甄选，也就是孟子所说的，以"权"知轻重，以"度"知长短。反过来讲，对人才的考试、考核、考察或甄选，就是检测和辨别人才德、学、才、识、体各方面的个别差异，获取恰当任使人才的科学依据。因此，考试在任何社会、任何国家的人事管理中，都是必不可少的手段。也正是由于考试在"了解人"方面有突出效用，历史上许多贤明的政治家、思想家才提出了"察能授官"、"居官必考"的主张，古代的各种选才、任官之法才为考试所取代，中国的考试制度才为世界众多国家借鉴与效仿。现在，考试已成为人事管理工作的根本要素，其功用十分广泛，大致可归纳为下列诸点：

（一）有效开发人力资源

在人事管理工作中，发现、识别人才，是合理利用人力资源的前提；培养人才，激励人才奋进，使之不断更新知识、提高素质、增长才能，是扩大人力资源的重要途径。抓好这两条，既是人事管理的基本职能，也是开发人力资源的关键所在。然而，“要研究人，要寻找能干的干部”①，要正确地履行以上职能，实现有效开发人力资源的目的，考试的功能实不可忽略。对人力资源的开发，考试的作用有三：一是通过考试可发掘人才，使社会上的“潜在人才资源”变为现实的人才资源。譬如，社会上有不少未接受系统的学校教育，在长期的实践中通过自学或其它途径成才的人，考试能给他们以表现知识、才能的机会，使那些确有真才实学的人获得相应的工作职位，尽其才能于社会；又如，对于各类待业人员中的德才兼备者，通过考试，若能达到既定标准，便可录用，让他们的聪明才智得到正常发挥；再如，部分学有专长、精力充沛的离退人员，也可经过考核聘为兼职工作者，让其“余热”贡献于祖国的建设事业。二是通过考试可发现人才，提高人才培养的有效性和社会人力资源的利用率。唐代诗人白居易在《放言五首》中写道：“试玉要烧三日满，辨材须待七年期。”人才的识别也是同理。在人才使用的过程中，只有经常地进行考察和考核，才能发现真正的人才。而对这样的人才进一步有计划、有针对性地培养，便是一种“贤明的利率最高的投资”，能使同一人才在相同的时间内作出更多的贡献，达到提高人才个体工作效率的目的。三是通过考试能不断扩大人力资源。科学的考试，常常是催人自新、激人奋进的动力。无论是资格考试、业务考核，还是竞赛选优、提职晋级考试，其考试结果对于干部、职工个人来说，不仅仅是反映已经具有的政治素质、业务水平和工作能力、或前期取得的成就，同时还反映自己知识的缺陷、能力的薄弱环节，以及本职工作的问题所在，使从业人员加深

---

①《列宁全集》第35卷，人民出版社1959年版，第542页。

自我认识，既了解自己的优点、长处，又明确自身的缺点、差距，并从比较分析中找出日后拼搏的方向，真正把生活的目的富于本职工作中，激励自己朝着新的目标努力奋进，不断提高自己的业务水平和工作能力，以适应日益变化的工作需求，推动事业的发展。此外，在某种程度上讲，考试是竞争的产物，而考试的结果又能促成新的竞争。客观准确的考试结果若能得到正确的使用，将有利于个人之间展开正常的比、学、赶、超活动，充分挖掘各自的潜能，进而形成新的人力资源。这样，即使从业人员的数量不变，也可大大提高单位工作效率。

（二）科学利用人力资源

准确评价人才，恰当使用人才，是人事管理的又一重要职能。它涉及人事管理工作的多方面，如对人才的评量、任用、晋升、提薪、调迁、奖惩等等，都与人力资源的利用密切相关。其中至关重要的是两条：量才准确，因才施用，真正达到“无能者不敢当职”，“有能者亦不得蔽隐”；劳酬相称，晋升、调迁有据，奖惩标准分明，做到“食有劳而禄有功，使有能而赏必行，罚必当”。①上述管理目标的实现，需有相应的制度、政策、法规给予保证，同时还必须有科学的考试手段与之配合。尽管考试结果的准确度是相对的，但考试终归有一个统一公认的标准，较之主观判定，它是当之无愧的“公正裁判”。在人事管理工作中，人们之所以普遍采用考试法，不仅因为考试能较为全面了解和掌握在职人员的情况，为各类人才的任用、晋职、调资、奖惩、调迁等提供真实的依据，更为重要的是，客观可靠的考试结果，对科学利用人力资源的功用显著。第一，由于考试对在职人员了解全面，能因才施用、各司其职、才尽其用，使一个部门或单位收到最佳整体效能。第二，通过考试能较公正地评价个人劳绩，有利于正确贯彻“按劳分配”的原则，做到劳与酬一致，能与级

①李悝：《说苑·政理》.

相称，功高者有赏，成绩低劣者有罚，贡献大者晋升，不称职者降级或调换工作职位。这对增强各类人才的工作责任感、调动积极性、鼓励先进创优、鞭策后进向上，扶持人才成长，都有积极的推动作用。第三，通过考试，用人标准明确，既有利于根除人才一次录用定终身，以及工资只升不降、职务只上不下的弊端，又有利于改变人才只进不出的僵化局面，可促进人才流通，加速各部门人才的新陈代谢，使不同水平和能力的人才都有用武之地。

（三）推动人事管理科学的发展

人事管理科学是一门颇富政策性、科学性、艺术性和实践性的科学，开展人事管理研究必须占有广泛的资料。这就需要有搜集素材、回收信息、理论与实践衔接的手段。而对各类在职人员进行多视角、多层次的考试、考核或考察所获得的结果，能比较真实地反映整个干部、职工队伍的思想状况、业务素质、适应能力、工作效率，它是人事管理研究必备的素材，人们可从中总结经验教训、探讨内在规律、使感性经验逐步升华为科学的理论。同时，目标、内容、方法不同的考试，能从不同方面及时反映人事管理工作面临的新任务、新问题，并暴露出人事管理理论上的缺陷，从而为人事管理科学的研究提出新的课题，促使理论研究工作者紧密联系实际，通过再认识、再探讨、深化、更新和发展人事管理科学理论，以指导和推动新的实践。

（四）促进人事管理制度的革新和完善

人事管理，实质上是对人才的一种控制。如果没有畅通的信息反馈渠道，管理就会运转不灵，甚至失控。在人事管理中，考试是获取反馈信息的重要手段。它能将各项制度、政策、法规、管理措施的执行情况，以及人才结构、素质、劳绩的现状与变化动态等方面的信息，反馈到人事管理部门，供管理者用作分析、研究的素材，并以此为依据，审查既定制度、方针、政策和法规

的科学性与可行性，再针对客观变化作出新的决策，使之充实和完善，不断增强人事管理制度主动适应社会发展需求的机制，始终保持人事管理部门的活力。

二、教育行政管理的必备措施

"考试在教育实施中要算是最古和最重要的了。它是学校工作之一，具有各种重要的功用，如决定教学方法的成败，供给成业的标准和激起优良的学校作业等等。它不仅是有用的必需的教室中的臂助，而且对教师的工作有密切的关系"，①同时，还具有辅助教育行政管理的作用，历来是实现教育行政管理目标的必备措施。它在教育行政管理方面的功能，可归结为检测、反馈、诊断三大功能。这些功能对教育行政管理的作用是：

（一）检测教育、教学质量和学业水平

教育行政管理的基本职能之一，是掌握教育动态、了解教学效果、检测学生学业水平、把好质量关。这一职能的履行，必须依靠考试的协助。比如，在一定时期内，全国或不同地区各级各类教育的质量，是升还是降？是全国性上升或下降，还是区域性上升或下降？是各类各级教育的普遍上升或下降，还是某类某级教育的质量上升或下降？与国外同一时期同类同级教育相比，质量是高还是低？教学成效是大有提高、略有提高、毫无提高或是降低？学生的学业水平是否达到各级培养标准？是全部学生达标，还是部分学生达标？是单科达标、部分学科达标，还是所有学科达标或全未达标？是教与学的原因，还是培养目标自身不科学？诸如此类的问题，都有赖于考试协助解决，即只有在全国或地区范围内，通过分科的或综合性的常模参考性考试，并对考试结果进行不同形式的比较分析，才能判定教育质量的升降及不同地

①〔美〕亚尔保德·兰著，浦漪人、黄明宗译：《新法考试》，正中书局1935年版，第1页。

区教育水平的差异，发现教学中的得失进退，确定教学成效的大小，找出影响教育与教学的因素；只有通过目标参考性考试，取得检定学业水平的客观依据，并参照各级教育的培养目标，才能对各级在校学生的学业成就作出科学的鉴定，并以此为据，给学生授予具有某级学业水平的证书。

（二）为教育对象分流、分类提供依据

学生因受生理和心理个别差异的制约，在接受教育的素质与能力方面，一般是因人而异，在同级教育中的适应能力和学业水平参差不齐，学术性向也不尽相同。为使教育符合人的身心发展规律，按照因材施教的原则组织教学，让不同素质、水平和学术性向的学生，都能在适合自己身心条件的环境中学习，教育行政管理部门及教师可借助考试提供的依据将教育对象分流或分类。当学生结束小学或中学的学业后，便可采用学能考试、学术性向考试，检测学生的学习能力，鉴别学生对不同类、科、专业教育的可接受性，以利教育行政管理部门对学生的甄选切合实际，使不同学习能力与学术性向的学生，分别进入不同类型的学校或大学的不同系、科、专业学习。对刚进校的学生，则可通过教学前考试（或称摸底考试）了解每个学生水平的差异，区分出每个学生的实际程度，进而为教育行政管理部门或教师恰当安置学生提供依据，把学生分别编入合适的教学班、组。此外，鉴别能力强的学术性向考试，还可较为准确地测定出学生的特殊才能，有利于特殊人才的培养和个性的发展。

（三）促进教学改革

从学校教育来说，考试与教学大纲、教学过程，是一种从属、交叉关系：教学大纲指导教学过程，考试寓于教学过程之中，而考试又对教学大纲和教学效果进行检测、验证，并通过反馈信息能动地作用于教学改革。在学校教育中，教学大纲的科学与否，需经教学实践检验，教学是否符合大纲要求，是否实现规

定的目标，单凭师生或教育行政部门的主观分析，是很难作出客观评价的，还需通过考试对教学质量进行检测，并对考试提供的各种数据进行全面的综合分析，才能得出科学的结论。同时，考试通过对教学成效的检测，又为教育行政管理部门和教学双方提供反馈信息，说明使用某种教材和教学方法后产生的实际效果，以及教学大纲和教学过程中存在的问题，帮助教育行政管理工作者、教师和学生查明问题的症结，以便选择有针对性的措施，或及时调整教学内容和教学方法，或修改课程目标和教学目标，使教学工作始终有效地进行，教学质量稳步上升。

（四）有助于了解教育行政管理的成效和改进管理工作

人们认为，学校的教育质量和教学效果好比"晴雨表"，能敏感地反映教育行政管理工作的质量。但是，要准确鉴别教育行政管理工作成效的大小，探明影响工作效率的弊端及其原因，达到改进管理工作的目的，首先需有检验工作成效、发现弊端的方法。标准成就考试、研究性考试或考核，就是最有效的方法之一。当某项教育方针、政策、措施、法令、规章制度、改革方案确立后，若要检测其有效性，并为进一步修改、完善取得依据，可在正式实施之前，通过标准成就考试或研究性考试，对学校各项工作、师资质量、教学质量等，进行一次全面的检测和评估。在实施的过程中或实施周期结束后，又以实施前采用的同样方法与内容，再进行一次全面性的检测。然后借助统计方法和电脑技术，对先后两次（或数次）考试所获得的各种数据资料作比较分析，可了解各项教育行政管理方案、措施、规章制度、政策、方针与方法实施前后的对比效应。若实施后的效果优于实施前，说明管理得法，切合实际，并通过对考试结果的分析研究，从中找出成功的因素，为进一步贯彻执行创设条件。若实施后成效不佳，甚至不如实施前，通过分析研究考试提供的资料和数据，也可查出失败的原因，有助于管理部门及时改进工作，或重新拟定方

案、制定措施、作出决策。

总之，教育行政管理、被管理以及他们共同作用的对象，之所以能够有机结合，除了有符合实际的定性要求外，更主要是因为有定量性的一致目标和为实现这种目标而建立的相互关系。如果说人们通常所讲的"胸中有数"就是某种管理目标的具体化，那么，"胸中有数"当是管理者科学管理的前提，胸中无数就很难在管理中做到适度，即实现质与量的统一，也很难准确检测管理工作的实效。因此，要检验教育行政管理工作的成效，单凭定性分析的结果是不足以为据的，还必须通过考试这一量化工具获取各种数量资料，进行定性、定量相结合的综合分析研究，所得结果，才能作为鉴定工作成效、改进管理方法与手段、探索管理规律的依据。

第三节 督导功能

在教学过程中，考试既反映教学的基本任务，又落实教学的基本任务，同时还督促教学双方完成教学任务，以及检验教学任务完成的质量。因此，考试对师生双方在教与学的目标选择和追求的动机方面，在实现抉择目标的意志力方面，都有极大的强化作用。它犹如一座桥梁，将教学双边活动有机衔接，好比一条纽带，把师生双方的力量携在一起，恰似一根指挥棒，能动地作用于教学双方的情绪与节奏，不断调整教与学的步伐，拨动教学发展的方向，使之步步趋向规定的教学、教育目标。如同日本梶田睿一博士在《现代教育评价论》一书中所指出："如能在教育中恰当地开展评价活动，并为教育工作而灵活运用评价的成果，那就有可能保证所有的孩子都具有一定的学力，有可能实现既不强迫又不浪费的恰当的教学指导，以及养成孩子对自己进行恰当的

自我评价的习惯，造成理想状态的教育。”也就是说，以考试为中心的评价措施及其结果的科学应用，对学校教育的成败有不可低估的督导作用。

一、实现教育目标的“整速器”

学校教育工作的主体是教学，其它各项工作都是为之服务的。如果学校不能圆满实现教学目标，也就不可能真正实现教育目标。从这种意义上讲，考试对于教育目标的实现具有“速整器”的功能，即调整教学步伐，强化目标意识，促使师生进取。考试是教学活动中的有机组成部分，本身负有培养人才的职能，它必须也能够体现教学和教育目标的内容与要求。而且，考试对教学达标度的检测不是一次性的终点检测，而是本着循序性、阶段性的教学原则，根据教学大纲的要求，将教学目标“细化”为若干递进性的阶段目标，并与教学阶段目标的节拍保持一致，使之成为整个教学进程不同阶段的“里程碑”。这样，考试对教学达标度的检测，既是阶段性的，又是经常性的；既是总结性的，也是起始性的。每一次考试，相对于前段教学是终点，而相对于后期教学则是起点，它始终督促师生朝着新的更高一级的目标迈进。

考试的“整速器”作用主要表现在两方面：

（一）强化教师追求教学目标的意识，督促教师根据教学目标的要求选择教学方法、调整教学内容，注重每一教学阶段的教学效果

教学过程中所举行的各种考试，都是为教学实现预定目标而设置的。每一次考试无不反映教学目标的一定内容和要求。除为合理安置新生而举行的学前考试外，不论是单元测验、期中考试、学期考试、学年考试，还是毕业考试、升学考试或论文答辩，其测试结果，总是包含两方面的内容：一是“学生掌握知识的数量与质量，技能的准确与熟练程度，以及能力的发展水

平”①；二是教师教学的效果。因此，通过考试结果，教师可了解前期教学的成效，检验教学是否符合教学目标的要求，深化教师对教学目标的理解，“并使教学目标对教师具有真实感和富有意义”，故有助于教师把教学目标寓于教学之中。而教师对考试所获反馈信息的分析和研究，实际上是教师对教学的一种自我反省，可鞭策教师进一步明确教学目标的要求，认真调整和改进教学内容与方法，准确把握教学的方向，提高教学效果。另一方面，教学过程中的考试多为教师亲自编制，他们参与确定考试目标、选择考试内容、编拟试题和拼配试卷的过程，也是检测教师掌握教材内容的程度，对教学目标深入了解及不断强化追求教学目标意识的过程。所以，“爱列亚脱判断哈佛大学教授的教学能力，系根据他们给与学生的各种考试，……考察许多试题后，就可知道教师的测验技术是否纯熟，对于教材是否了解。”②美国测验专家亚尔保德·兰也认为：“教师的‘教学方法与实际’的知识可用专为那种目的而编的职业测验以测量之。”③

（二）强化学生的学习动机，激发学习兴趣，促进智能发展

对在校学生来说，考试是一种压力，但更是一种动力。它“可以促进学生及时复习功课，巩固和加深知识，培养精确细致、刻苦认真的学习态度，系统的劳动习惯和坚强的意志品质。”④既是强化动机培养进取精神的有效措施，也是教学中促进学生智能发展必不可少的手段。

大量的实验研究表明，考试对学生学习的影响与作用，在考

①潘懋元主编：《高等教育学》，人民教育出版社和福建教育出版社1984年版，第295页。

②③〔美〕亚尔保德·兰著，浦漪人，黄明宗译：《新法考试》，正中书局1935年版，第25页。

④潘懋元主编：《高等教育学》，人民教育出版社和福建教育出版社1984年版，第296页。

试前后两个阶段均有极为明显的反映。如果在实施考试之前，教师能以恰当的方法通知学生准备应考，并给予科学的指导，使学生形成良好的备考意识和应考心理，将会收到较之平日更高的学习效果：首先，考试可使学生的思维高度集中，实现增强学习意志、提高学习效率的目的。绝大多数学生为获得好的考试结果，考试的压力能转化为勤学苦钻的动力，增强学习的意愿，把时间和精力集中于某一学科或某几学科，从而"导致大量的学习"。单位学习的效率也因此提高。其次，由于备考为学生创设了"再学习"的条件，对已学内容至少增加两次重新学习、钻研的机会，即分别通过备考和考试试题的解答重温、反思前期功课。这种"再学习"并非是简单的再认过程，而是对所学知识和技能巩固深化的过程，有助于学生增加学后保存量，以及对已学内容的消化、理解，进一步掌握所学知识之间的内在联系，收到触类旁通之效，也有利于学习迁移和学生自学、思维等能力的发展。第三，如果考试的方法、内容科学，能真实反映学生掌握知识的质与量、技能的准确熟练程度和已经具备的能力水平，并将测试结果量化为准确的考试分数，使之成为体现学生学业成绩差异的客观标志，考试就会对日后的学习产生积极的推动作用。它能够给学生"提供一种熟悉而又相当稳定的形式，使人能将生活目的寓于其中"①，向新的学习目标进取。相反，倘若"没有评分，学生的学习便缺乏引导，而且也丧失了追求一定目的意识"②，削弱学习的动机与兴趣。还应指出的是，参加考试的过程，也是学生加强自我认识的过程。他们从考试反馈信息中，能了解学习的成效，领会到运用所学知识、技能服务社会的意义，以及知识、技能的实际价值，这对学生的学习动机与进取精神，是一种激励和强化。再者，考试成绩是多种因素相互作用的结果，它提

①②〔美〕罗伯特·蒙哥马利著　黄鸣译：《考试的新探索》，广西人民出版社1984年版，第26页。

供给学生的反馈信息，也不只是知识、技能和能力方面的内容，同时还在一定程度上反映了学习目的、学习态度、学习方法、学习能力、意志品质、身体素质等多方面的情况。因此，考试可帮助学生更好地了解自己，对学习中的优点和缺点、自身的长处和短处，有一个比较全面客观的认识，促使学生在新的学习中扬长克短，不断提高学习的效果。

二、后继人才智能发展的“导标”

在我国历史上，人们素有“科举有甚于十万督学之力”的说法。这话深刻揭示了考试对师生教与学的督促作用。到了近现代，随着人们对考试功能认识的深化与发展，认为考试对教学双方还有引导的作用，被人们喻之为“向导”、“指挥棒”。现在，考试对教学的引导作用虽为社会所公认，但对考试引导作用的看法仍有偏见，比较突出的是：多数人对考试的引导作用并不是正面肯定，而是一种带有责备性的评价，甚至作为现行考试的主要弊端进行批驳。这种看法是不全面的，它仅仅是从考试某些不甚科学的因素对教育产生的不良作用方面得出的结论，而忽略了考试对学生智能发展积极引导的一面。另外，考试的引导作用并不局限于“它对教师应如何教、学生应如何学的左右和影响”，而是教育内外诸方面，关系到整个后继人才的成长。

考试对后继人才智能的发展具有导航的作用，这是考试在人才培养措施中的特殊地位所决定的，也是考试教育性的必然反映。这是我们在选择考试的内容与方法前必须想到的。“对于功用如不加以考虑，就是最有价值的活动，也不会产生最大的效率，它们的各种可能性也只能获得部分的利用而已。”①甚至走向反面，正象沈起炜在《说考》中所比喻的那样：“考试好象是

①〔美〕亚尔保德·兰著，浦漪人、黄明宗译：《新法考试》，正中书局1935年版，第15页。

水、火。人没有水、火，无法生活，……然而水、火也能成为人的大祸。”就是说，考试的“导标”作用具有二重性：科学的考试可给后继人才的智能发展导以正确方向，引入健康发展的轨道；若考试内容方法不当，又将压抑后继人才智能的发展，或将智能发展导向偏枯。其原因有二：

一方面，处于求学期的后继人才，其智能发展的外部主要因素是教学内容和教学方法，而考试作为教学活动中的一个组成部分，一种育人的重要手段，它在很大程度上支配着师生教与学的步骤和意识倾向，尤其在学生对学习内容、方法、目标、乃至发展方向等的选择方面，考试具有导航的作用。

“考试引导学生学习，往往是前一阶段的考试，引导着学生后一阶段学习，即使教学内容已经转化，学生仍然可以从考题中接受引导，从以前的轨迹中找到今后的门径。”①至于考试能否将学生的学习引向正确的门径，能否将后继人才的智能发展导入健康轨道，则取决于考试本身是否科学。只要考试全面体现教学过程的基本任务，如实反映教学的内容与要求，并有科学的方法、手段与之配合考试，就能正确发挥教育作用，引导后继人才向着学校教育预期的方向发展，形成与其社会所期望的智能结构。如若考试脱离教学实际，随意确立标准和拼凑内容，方法不当，也难免会出现清人严几道在《救亡决论》中斥责“八股”考试时所指出的“三大害”：“曰锢智慧，曰坏心术，曰滋游手”。长期以来，考试实践的客观效果也说明，缺乏科学性、不符合个体智能发展规律的考试，确有将后继人才的智能发展导向偏枯的不良作用。比如：考试内容偏重知识、技能的再认和再现，就会束缚后继人才观察、想象、思维等能力的发展；考试方法限于笔试，就会影响后继人才口头表达、实际操作能力的发展；考试题型单

①张大展等：《学校在学考试作用初探以及改革中学在学考试设想》，载《福建教育科研通讯》第9期，第9页。

一、内容过窄、应答方式死板，后继人才独立获取知识、创造思维等能力的发展，都会受到程度不同的约制。

另一方面，后继人才的求学期，也是他们身心发展的关键期，正由不成熟步步走向成熟，渐渐形成具有个性倾向和个性特征的心理系统。这种心理系统本身含有许多矛盾性和新旧力量的对立因素，从而构成了后继人才各自内心世界的矛盾性和力量对比。而在个性心理系统的诸种新旧对立因素中，又以新的愿望和需要最为积极、活跃，每个人总是以满足自己新的需要和实现新的愿望为目的，对环境和教育作出相应的选择，确立新的心理因素发展的方向。在此期间，影响后继人才智能发展的因素，除自然、社会、家庭环境，教师的品质，教学的指导思想，教学内容和教学方法外，还有考试这一重要因素。由于在校考试成绩的优劣标志着学生成就的大小，考试成绩达标，在某种程度上意味着社会的承认，若是毕业或升学考试，其成绩还与后继人才未来的职业、经济待遇、社会地位等密切相关，所以，考试最容易成为左右后继人才智能发展方向的"导标"。它对后继人才智能的发展也许是扶持，为其导之以正确方向；也许是压抑，导致后继人才的智能畸形发展。只有使考试真正成为被试者个性心理系统内在矛盾新旧力量对比发生积极变化的外部有利条件，才能产生引导后继人才智能健康发展的效果。

还应着重指出的是，考试这种测试人的德、学、才、识、体个别差异的活动，之所以为人类社会不可缺少，就在于它具有多种功能，不失为人才培养、选拔、任用、管理的一种良方，如果看轻考试的作用，或随意废弃考试，都将于国于民不利，就象法国总统蓬皮杜所说的那样，"会是一国之灾难"。但是考试也并非万能，如果任意夸大考试的作用，以为考试能指挥各行各业，或可替代其它各种育才、选才以及人力资源开发与外用的手段，随意滥用考试，无疑是违背科学的，同样会贻害人才，断送事业。

综上所述，考试作为人类社会母系统中的一个子系统，它受制于多方面同时又作用于多方面，它的发展有赖于母系统的推动，同时又是推动母系统发展的重要机制，它的功能是多元的，同时又是一个互有关联的整体。因此，我们若想正确发挥考试的功能，使之有效服务于社会，首先必须科学分析考试与其它子系统的关系，探明考试各种功能之间的联系与区别，把握考试功能的二重性，对考试功能有一个全面的认识。

第四章

考试与人才

人才是一种历史概念，考试是一个历史的发展过程，两者互相制约、相互促进，既有历史的继承性，又有明显的时代特征。

人类社会演变发展的史实表明，人才既是创造物质文明和精神文明的中坚，也是振兴民族、治国安邦之本，恰如墨子所云："国有贤良之士众，则国家之治厚；贤良之士寡，则国家之治薄。"①对此，列宁也有深刻的论证，他在论及苏维埃政权的任务时明确指出："没有具备各种知识、技术和经验的专家来指导，便不能过渡到社会主义。"②所以，人才，"国之宝也。"③而人才与考试的关系是十分密切的，甚至可以说，只要人类社会存在选人、用人问题，就必然有测度、甄别人之素质与智能优差的考试存在。

在这一章中，我们将通过对人类人才观、考试观演变史实的回顾，对考试与人才的关系作较为详细的探讨。此外，对人才的定义、人才的类别和人才质量的基本标准问题，也将进行阐述。

①《墨子·尚贤》

②《列宁选集》第3卷，人民出版社1972年版，第501页，

③《明史·选举志三》

第一节 人才观、考试观的演变与关联

随着人类社会的发展与演变，人才观和考试观也相应历经了产生、发展、形成和演变的过程。那么，这个过程的情况究竟怎样呢？正是本节所要讨论的问题。

一、人才观的演变

由于人类征服自然、改造社会与自身发展的客观需要，自古以来，人们始终在探讨一切社会所面临的一个共性问题：如何有效地考察、甄选、任使和管理人才。正是在这种不间断的实践和研究中，渐渐形成了人类社会最早的人才观和用人原则，并随着这种实践的发展和认识的深化，人类的人才观和由其决定的用人原则，也渐趋成熟和科学。

（一）天下为公的用人原则

原始社会最初遇到的重大人才问题，大概是挑选部落（氏族）的首领。首领产生的形式和方法是由部落（民族）的成员推举或在职首领选定。据古老的传说和我国最早的文献记载，我们中华民族早期的一些首领如神农、黄帝、燧人氏、尧、舜、禹等，都是凭自己的智慧、经验、胆略和勤奋为民族作出了贡献的人。如神农教民耕作，发展了农业，为了给人们治病，遍尝百草；燧人氏为了使人民吃到熟食，发明了钻木取火之法；尧身为天下之首领，仍住茅棚、吃粗米、食野菜，与民同甘共苦；禹为根治洪水患之，带领民众，身先士卒，三过家门而不入，历尽了千辛万苦。由此看来，原始社会的用人原则是：天下为公。只要有才能，并能真心实意地为大家办事的人，即使其父母犯了严重过错，也不会妨碍大家对他的信任，禹获得舜的推荐并禅位，便是

一例。

原始社会后期这种“天下为公”、“唯才是举的选人、用人原则，之所以能产生并延续到奴隶社会初期，一方面是由于当时生存环境的恶劣和人类生存能力的低微，为求得生存和发展，人们只能也必须这样做。另一方面，在那种“原始共产主义”的境况下，部落（氏族）的首领也只有全心全意地为全体成员谋生存、求发展，作名符其实的社会“公仆”，才能赢得众人的信任和拥护。也就是说，“天下为公”、“唯才是举”的用人原则，虽然是自发的，但也是这一时期社会环境和生产能力的必然结果。

值得特别指出的是，正是原始社会中人们为了求生存、求发展而自发形成的这种“天下为公”、“唯才是举”的用人原则，对我国几千年的人才研究和中华民族人才观的形成与发展，以至对全人类人才观的形成与发展，都产生了极其深远的影响，即使在当今时代，也仍然具有十分积极的意义。

（二）君主型人才观向官吏型人才观的演变

从国家的出现到英国的资产阶级革命以前这一历史时期，被史学家们称之为奴隶社会和封建社会时期（在各不同的国家中，封建社会结束的时间有迟有早，如我国，封建社会一直延续到本世纪初）。由于这几千年的生产活动，基本上是以农业、手工业为主体，社会经济的基本特征是自给自足的自然经济。从此种角度，我们将人类社会这一历史时期称之为农业、手工业社会。

在农业手工业社会的第一阶段，即由原始公社（原始部落、原始氏族经济）转变为畜牧业、农业经济的奴隶社会时期，由于原始社会中部落（氏族）“首领”的形象和影响尚存，奴隶国家的统治者——奴隶主们出于巩固统治的需要，在加强对过去首领的神化宣传，借以烘托和树立自己形象的过程中，对如何才能当好“人君”也展开了积极探索和研究，进而形成了这一时期具有代表性的人才观——君主型人才观。

君主型人才观认为人才之德是君主之德，人才之才是君主之才。也就是说，这个时期人才德与才的标准，主要是以什么样的君主才是好君主而具体化的。关于君主人才观的德才标准，尤以我国先秦时期的人才思想家们研究最深，其论述也最具代表性。例如《管子》一书，就对“君主人才观”的德与才作了明确的阐述。他认为，君主之德为“爱民”：“民恶扰劳，我逸乐之；民恶贫贱，我富贵之；民恶危坠，我存安之；民恶灭绝，我生育之。”而君主之才则是：善于教化——以道化民，教子学技；善于用人——问其乡里，参其成事，设问国患；察能授官，任其所长；爵授有德，禄与有功；他酬一倍，我酬三倍；善于防奸——建立法令，蔽贤受罚；毁誉之人，勿与任大；等等。①

在奴隶社会向封建社会过渡的变革中，西方的思想界未曾出现人才辩论的鼎盛时期，而在我国却出现了历史上第一次关于人才问题的大辩论，造就了一大批社会变革急需的人才。老子、孔子、墨子、商鞅、孟子、荀子、晏子和韩非子等就是其杰出的代表。

在这场空前的思想、学术的“百家争鸣”中，人们就人才的标准、人才的重要性和人才的培养、使用等一系列问题展开了激烈的争论，从而大大推动了当时人才理论的研究和发展，加速了君主型人才观向官吏型人才观的演变。

封建社会的欧洲处于野蛮的中世纪，而与欧洲的中世纪相对应的我国的春秋战国至汉唐时期，却正是封建文明兴盛之际，所以我们仍以我国这一时期的情况为例，对“官吏型人才观”的特征作进一步探讨。

与君主型人才观相比较，官吏型人才观有以下新的特点：第一，重视“官吏”在治理国家中的作用，主张“任人唯贤”“察

---

①参见《管子·霸言》。

能予官，以德就列”①。第二，“人才”的范围扩大了：君主是人才，官吏是人才，凡有一技之长者也是人才。如善医者扁鹊和善相马者伯乐，均被列入人才之列。第三，主张不拘一格选用人才：“官无常贵而民无终贱”②，“农与工肆之人，有能则举之”③。第四，关于人才的内涵拓宽了：凡能“为之于其未有”者④，都视为“才”。如射箭本领高超的羿、医术超群的扁鹊、烹调技术高明的庖丁、抚琴能手伯牙、精于相马术的伯乐，等等。

官吏型人才观不仅对人才的德提出了新的内容，而且明确提出了德与才之间的关系。在君主型人才观中，人才的德就是爱民，而官吏型人才观中则演变为“忠君爱民”。其具体的内容就是封建的忠、孝、礼、义。在人才的德才关系方面，则明确提出“才者，德之资也；德者，才之帅也”⑤。也就是说，德靠才来发挥，才靠德来统帅。

人类的人才观从“养人如养六畜，用人如用草木”⑥的君主型人才观演变为视人才若“国家之珍”的官吏型人才观，是人类人才观的一大进步，也是人类人才认识史上的一次飞跃。

（三）官吏型人才观向专业型人才观的演变

欧洲的文艺复兴加速了中世纪野蛮、黑暗封建制度的解体，也加快了资本主义生产发展的步伐。同时，在“向生命索取一切它所能提供的东西”口号的鼓舞下，觉醒了的欧洲在科学、文化、艺术、教育……几乎一切方面都表现出了极大的热情。“这是一次人类从来没有经历过的最伟大的、进步的变革，是一个需要巨人而且产生了巨人——在思维能力、热情和性格方面，在多才多艺和学识渊博方面的巨人的时代”。⑦

①②③《墨子·尚贤》

④《老子》第六十四章。

⑤司马光：《资治通鉴》。

⑥《管子·七法》。

⑦《马克思恩格斯选集》第3卷、人民出版社1972年版、第445页。

文艺复兴——技术、经济和社会（文化、教育、艺术等）三方面联合起来的革命，既使资产阶级登上了社会的政治舞台，也促进了第一次工业革命的出现。于是，一种崭新的带有几分神奇色彩的生产工具——机械出现了，并迅速地被应用于广泛的生产领域。随后，由于科学的介入，又出现了第二次工业革命。第一次工业革命的产物——机械，与第二次工业革命的产物——电能的结合，它不仅给人类社会的生产力带来了几十倍、成百倍的增长，而且强化了社会各行各业的竞争。社会对人才数量的需求迅猛增加，对人才的质量提出了崭新的要求，于是，一种反映机械化大生产时期社会需求的新型人才观——专业型人才观取代了农业手工业生产时期的官吏型人才观。

与官吏型人才观相比，专业型人才观的特点是：第一，行行要人才，行行出人才，行行有人才。第二，强调人才之才的专门性，是专业型人才观最根本的特点。所谓人才之才的专门性，是说人才者应该精通某一门学科，或某一种技术，而且越精通越好，至于其他方面懂不懂或懂多少都无关紧要，所强调的是"专才"，是一技之长。第三，承认人才是有层次的。在纵的方面，社会的各个行业的各个层次都有人才；在横的方面，社会的每一行业、每一层次中的人才，又有高、中、初之分。最后，这种人才观把人才、社会和教育视为一个有机的整体。从我国新民主主义革命的先行者孙中山先生的人才思想中，就可以看到这种专业型人才观的反映。他在《上李鸿章书》中提出："用人也，务取所长之其职。"①"凡学堂课此一业，则国家有此一官，幼而学者即壮之所行，其学而优者则能仕，且恪守一途，有升迁而无更调。"②这样，"久任则阅历深，习惯则智巧出，加之厚其养廉，永其俸禄，则无瞻顾之心，而能专一其志。"③不难看出，孙中山先生反复

①②③孙中山：《上李鸿章书》，见《孙中山选集》，中华书局1964年版，第2～3页。

强调的正是人才之才的专门性和如何正确使用这种专业型人才。

至于人才的德，专业型人才观所要求的是：平等、自由、博爱和爱国。必须指出的是，这种人才观的德的具体内容，在不同社会制度的国家中是不一样的。譬如在我国，解放以来关于人才的德的基本标准就是：热爱祖国，热爱人民，热爱社会主义制度，拥护党的领导。而在资产阶级统治的国家中，关于人才的德的标准就是另一码事了。

（四）专业型人才观向通才型人才观的演变

电子计算机的出现，以及它在生产、生活和科学领域中广泛的推广与应用，把人类社会推进到一个崭新的历史时代——电子时代。人类在这个时代才生活了40年，可谓刚刚开始。这一崭新时代的基本特点是：

其一，知识正以惊人的速度激增。据科学学家统计，如果把1750年人类知识量作为2倍，那么到1900年人类的知识量为4倍，1950年为8倍，1960年为16倍。目前，世界上平均每天发表1.3～1.4万篇论文，各种书籍每年增加25万种，每天有800～900项专利问世！

其二，知识陈旧的周期正在缩短。据统计，18世纪时知识陈旧的周期为80～90年，19世纪到20世纪初为30年，近50年中又缩短为15年，如今，一些学科领域中的知识陈旧周期已缩短为5～10年。

其三，科学技术由潜在的生产力转变为直接的生产力的过程越来越短。表4—1中列举的一些重大发明转变成直接生产力的情况，便足以证明这一点。

其四，物质生产的产品更新换代愈来愈快。如电子计算机平均每8年体积缩小10倍，成本降低10倍，运算速度提高10倍。

其五，新的技术革命正方兴未艾。它包括微电子技术、生物

表4—1　近现代部分重大科学发明（发现）转化为生产力时间情况表①

发明（发现）项目	发明（发现）年　份	出产品年份	发明（发现）到投产间隔（年）
摄影术	1727	1839	112
电动机	1821	1886	65
电话	1820	1876	56
无线电	1867	1902	35
真空管	1884	1915	29
X光管	1895	1913	18
雷达	1925	1940	15
电视	1922	1934	12
核反应	1932	1942	10
原子弹	1939	1945	6
晶体管	1948	1951	3
太阳能电池	1953	1955	2
激光器	1960	1960	1

工程技术、航天技术、遗传科学、材料科学、能源科学等。这场革命的结果，不仅将使传统的生产方式面目全非，并且对人们的生活方式、思维方式等，也势必带来深刻的变化。实际上这种变化已经开始。

如果以生产形式的特征来命名或划分历史时期，根据近40年来的变化及其发展趋势，我们完全有理由将这个崭新的时期命名为自动化生产时期。

科学技术日新月异的发展，生产、生活领域的急剧变化，对人才提出了不同于以往任何一个时代的要求。关于自动化生产时期的人才应该具备什么条件的问题，人们已经讨论、研究了至少30年。在这30余年讨论、争论的过程中，一种新的人才观——通

①《学习生存——世界教育的今天和明天》，上海译文出版社1981年版，第127页。

才型人才观问世了。虽然这种新人才观中的某些观点尚在争论之中，界定尚不成熟，但它的一些基本观点，则比较正确地反映了自动化生产时期社会对人才的基本要求。这些观点是：

第一，自动化生产时期的人才应该是智能型的，而不应只是知识型的。具体地说，新时期的人才应该具备五种能力：自学能力、知识更新能力、研究能力、表达能力和组识管理的能力；要有宽厚的理论基础，较强的适应能力和应变能力；既有顽强的进取精神和强烈的竞争意识，又善于与他人合作；要有强壮的体魄和优良的体能；要有良好的美感和审美能力等。

第二，为了保证人才的质量，传统的教育必须改革。学校，特别是高等学校教给学生的不应只是“定型的知识”，而且应使学生“学会学习”①的“能力”。

第三，人才的真正价值在于使用。对人才的管理应该是动态的，让人才在流动中施展才能、发挥专长。

第四，人才的才能是有等级的，对人才的使用应该坚持“能级原理”，真正做到才尽其用。

第五，人才的培养、使用和管理的全过程，是一个有机的系统。将这个过程中的任何环节割裂开来、对立起来的作法，都是不科学的，因而也是有害的。

自动化生产时期的人才观的最终形成，将有待于这个时期的进一步发展和它对人才的要求的进一步明朗化和相对定型，也有待人们、特别是人才学家们更加深入系统的研究。

总之，从人类原始社会自发的“天下为公”、“唯才是举”的用人原则的问世，到通才型人才观的兴起，人类的人才观经历了一个漫长的演变过程。随着社会的发展，这个过程还将持续、延伸。

---

①德拉高尔朱布·纳伊曼著，令华等译：《世界高等教育的探讨》，教育科学出版社1982年版，第63页。

二、考试观的演变

由人才观派生并为其所支配的考试观，也有其产生、发展的漫长演变历史。

（一）考试观的萌芽

在原始社会的后期，部落或氏族首领的产生，最初是由大家推选，后来发展为既可由大家推选也可由在职首领推荐。但无论是推选还是推荐，都有一个选择什么样的人和用什么方法选的问题。前一个问题的实质是用人标准，后一个问题则属考试的方法和形式的范畴。人们在推选部落或氏族首领等活动中，逐渐形成的选什么人与如何选人的观念，便是人类考试观的萌芽。

原始社会的人才甄选，特别是部落或氏族首领的推选，实际上有一个德与才的标准问题。但在当时人们的头脑里，还不可能有明晰的德与才的概念，只能是根据谁能使部落（或氏族）生存下去并得到发展就推选谁。显然，这里面既有德的要求，也包含了才的条件。同时，要从人群中推选出符合要求的首领，还需有甄别人才素质与智能优劣的途径和方法。在当时的历史条件下，解决这个问题，只能通过观察日常生活及生产中的实际表现来鉴别。当绝大多数成员都认为某人的“办法”多，并愿意为大家服务时，这个人就很自然地（当然也有一定的认可程序和仪式）被大家推举为首领。这种以观察日常生活、生产的实际行为来鉴别、选拔人才的方法，就是操演测试法的肇端，同时也是当时德、才二者融为一体的一元型考试观的必然产物。

（二）重德轻才型考试观向分类统一型考试观的演变

随着原始部落（氏族）的解体和奴隶制的形成，原始社会时期的一元型考试观演变为奴隶社会的重德轻才型考试观。到了封建社会，重德轻才型考试观又逐步演变为分类统一型考试观。

上述人类考试观演变的历史轨迹，在我国上古至中古时代的

人才甄选中，体现得至为明显。在春秋末期以前的奴隶制社会中，国家的高级官员基本上是世袭的(改朝换代时的情况例外)。周朝虽设有国学和乡学，但由于只有进国学者才可能成为高级官员，而国学又只收贵族子弟，所以高级官员实质上是世袭。只有低级官吏是通过选士的方法产生和补充的。据文献记载，当时选士所测试的内容，主要是关于德行方面的要求。如选士的形式之一的学校选，其考试内容为“六德”、“六行”和“六艺”。选士的另一种形式是乡里选。乡里选不通过正规的考试，被选人只要得到由下而上的逐级推荐，即可入官。而关于征询乡里舆论的方式，则是在乡府典礼上，以“五物询众庶，一曰和，二曰容，三曰主皮，四曰和容，五曰兴舞”①。如果众庶一致认为被选举者具有以上五种品德才能，并且能“布之教于四方，父义、母慈、弟恭、子孝、内平、外成”②，就可以“出使长之，入使活之”③。由上可知，我国春秋时期考试观的整体特征是重德轻才的。这种重德轻才型考试观的特点，在这一时期的官员考察中同样有明显的反映。如周朝考察在职官员的“六计”，即“廉善、廉能、廉敬、廉正、廉法、廉办”，就是置德于考察的首位。

到了战国时期，由于社会生产力的发展和新兴地主阶级的兴起，社会政治制度发生急剧变化，从而打破了王室独尊、诸侯国林立的局面，几百个小封国归并为几个相应的大国，我国开始进入封建社会。此时，各国的新兴地主阶级出于巩固封建统治的需要，都十分重视招纳、选用贤能之士。于是，内正法、养士制度和游说取士等取代了周王朝的选士制度，并成为甄选和招揽人才的主要形式。其选用人才的突出特点，是重视人的才能。如在孟尝君三千食客中，不仅有智、勇、辩、力之士，甚至有善鸡鸣狗盗之辈；张仪、苏秦徒步为相；白起、王翦能成为名将，都是有

---

①②③《周礼·地官》.

力的证明。当然，无论是内正法，养士制度还是游说取士，虽注重人之才能，但并不等于对人才的选用没有德的要求。

战国时期选用人才虽看重才能，但对人才之才能的甄别尚无分类之举。这与当时各门科学尚无明确分类是相一致的。由于这一时期的考试观既重视人才的才，也重视人才的德，而对人才的德与才的关系又尚无明确区分。所以，我们将这种视德、才同样重要的考试观，称为德才并重型考试观。

在人类考试思想的发展史上，战国时期德才并重的考试观占有重要地位，它较之春秋时期重德轻才的考试观是一大进步。也正是在这种考试观的支配下，当时的人才甄选才得以满足社会的人才需求，推动了当时社会的进步。

德才并重型考试观始于战国，完善于两汉。在其思想的影响下，相继建立了“察举制”、“制举制”、“举逸才”、“九品中正制”，以及对官吏的“上计”、“刺察”制度等。但由于这些甄选人才的制度都是由各级官吏推荐，而并非正规的考试，且对人才的才能未制定客观的考核标准，加上当时社会政治制度固有弊病的影响，这一时期无论是“察举制”、“制举制”，还是“九品中正制”，实际上成了巩固门阀势力的工具。达官权贵们所推举或选用的人才，大都是与自己沾亲带故的人，造成了“上品无寒门，下品无势族”的现象，严重阻碍了社会的发展。

到了公元7世纪初，隋文帝开科举制之先河，并为后世沿用1000余年，成为隋至清末历代测度、甄别、选拔人才和官吏的基本形式。

随着选拔人才的考试形式和方法的变革，特别是这一时期人才理论的发展和人才观的演变，人们的考试观也发生了显著的变化，它由前一时期的德才并重型考试观演变为分类统一型考试观。其基本特点是：

其一，把选用官吏的权力收归朝廷，使考试为巩固中央集权

的政治制度服务。

其二，对人才知识才能的测试，有了统一明确的标准。

其三，德才关系有了明确的划分："才者，德之资也；德者，才之帅也"。对人才德的考核具有共同标准，而在知识、才能方面则分类测试。这有利于对不同知识才能的人择其长而用之。

其四，对人才的甄选，在一定程度上体现了以考试成绩优劣为据的原则，这对调动社会中、下层办学的积极性，促使读书人勤奋学习，起到了某些积极作用。

（三）德帅才型考试观向专才型考试观的演变

由技术进步引起的第一次工业革命和因科学发展而导致的第二次工业革命，使欧洲各国的政治结构、经济结构、技术结构、产业结构和人们的知识结构发生巨大的变革。随着这种变化而出现的对人才素质与智能要求的变化，又导致了近代考试的出现。而近代考试从目的、形式、方法到内容较之从前的一切考试都有显著的不同：

首先，考试的目的变一元为多元，即施考的目的由过去单一的选官取士，转向为各行各业（行政官员也是一种职业）选拔合适的人才、检测教学效果、评估教学质量、检验工作成就、考核业务技术水平等各种目的施考。

其次，考试的形式和种类多样化。可以这样说，社会上有什么职业，需要什么样的人才，就有什么样的考试；学校设有什么课程，就有什么考试。

再次，更为注重选才、育才、用才、管才之间的关系。

最后，考试逐步向系列化、系统化、标准化发展。

近代考试目的、形式、种类、方法和内容的变化，正是近代考试观——专才型考试观的具体反映。这种考试观的核心是：在坚持德帅才的原则下，一切考试都要为了培养、选拔和任用专门

人才服务。

（四）专才型考试观向系统型考试观的演变

第二次世界大战以来，科学技术的发展一日千里，并产生了前所未有的新特点：第一，既高度分化又高度综合，出现了一系列的系统科学，如空间科学、生命科学、海洋科学、环境科学、材料科学、能源科学、决策科学、未来学等等；第二，自然科学与社会科学合流，科学一体化的进程十分迅速；第三，科学和技术一体化，在诸如激光、电子、材料等领域内，孰为科学孰为技术几乎无法区分。科学技术的这些新特点，加上整个社会科学化进程的快速发展，现代社会对人才提出了更新的要求。（详见本节"考试观的演变"）

随着社会对人才要求的变化，人们的考试观也正在出现新的演变，即由专才型考试向系统型考试观演变。这种演变情况，工业化国家出现于50年代末、60年代初，我国则晚一些，直到70年代后期才开始。伴随着人才观、考试观的演变，现代考试的形式、方法和内容也出现了明显的变化。现代考试不仅重视量尺标准、实施规范，而且要求方法科学、手段先进、结果可靠和社会效应良好，能对社会的发展起到积极的推动作用。

系统型考试观虽然是历史上各种考试观演变发展的结果，但与以往的各种考试观相比，又有其鲜明的时代特征：

第一，考试不再是简单的选拔人才的工具或手段，而是一门对人才的培养、选拔、使用和管理具有特殊意义的科学。

第二，人才的标准是德、学、才、识、体兼备，考试不仅要全面检测人才德、学、才、识、体的实际水平，而且要促进人才素质与智能的全面发展。

第三，考试除了为人才的评鉴、选拔、任使、晋升、革职、调迁提供依据外，还合理调节社会各部门的人才结构，以及各类人才的知识、智能结构。

第四，考试与教育、人事管理有机结合，教、考、用三者间配合密切，运转协调。

三、考试观与人才观的关系

从本节上述的讨论中，我们已不难看出，考试观与人才观之间有着极为密切的关联。

（一）人才观决定考试观，考试观体现人才观

这是人才观与考试观之间最本质因而也是最重要的关系。

原始社会自发的、"天下为公"的用人原则，决定了当时德、才融为一体的一元型考试观。部落（氏族）的成员（包括在职首领）通过对人才生活与生产实际行为的考察，从整体上鉴别人才的德、才优劣，选择部落或民族的首领。封建社会时期的官吏型人才观，对人才德的要求是忠君、孝廉、克己、奉公，而对才的要求则有所侧重，即因官吏类别的不同而有别。这一时期的考试观为统一分类型考试观，即分科取士，择优授官。现代社会的通才型人才观，视人才为国家、民族及至全人类的宝贵财富，并将人才的培养、选拔、使用和管理视为一个有机的密不可分的系统，认为人才应该以德为帅，德、学、才、识、体全面发展。只要我们稍微研究、考察一下现代的考试理论和考试方法，特别是在目前教育改革中出现的一些新的考试思想、考试理论和考试方法，就会发现，它们所努力体现的正是这种通才型人才观。传统考试思想、考试理论和考试方法中那些不能体现通才型人才观的部分，正在被革新和改造。也就是说，目前世界各国改革考试思想、考试理论和考试方法的热潮的出现，是现代通才型人才观在考试领域的必然反映。

（二）考试观的形成和演变，决定于并滞后于人才观的形成和演变

考试观是随人才观的变化发展而变化发展的，它必须如实反

映人才观的要求。所以，考试观的形成和演变决定于且滞后于人才观的形成和演变。正确认识考试观与人才观之间的关系，其意义是多方面的：

在实践方面，它为我们评价一切现有的考试思想、考试理论和考试方法，提出新的考试思想，探讨新的考试理论和模式，提供了一条最根本也是最重要的原则：是否有利于当时社会所需人才的培养、选拔、任使和管理。因为任何历史阶段的人才观总是反映当时社会对人才的需求，所以，凡能较好地反映同一历史时期人才观的考试思想、考试理论和方法，就是正确的、好的，就应给予肯定。反之亦然。这一原则，对我们确定和评价具体的考试内容，同样具有指导意义。

在理论方面，正确把握考试观与人才观的关系，有助于我们全面认识和正确揭示考试发生、发展及演变的规律，使我们能够随着社会发展对人才需求的变化，自觉、主动、及时地探讨、革新考试思想、考试理论、考试内容和考试方法，使之与社会发展的人才需求保持一致，更好地服务于不同社会的政治、经济、文化、科技及其他各个领域。

第二节　人才的标准及分类

考试的全部意义和价值，在于为社会发展甄别、选拔各种人才和对人才的培养情况、人才的管理作出评价。具备哪些条件的人才是人才，培养中的人才是否达到预定目标，怎样才能确定某一人才或某一些人才适合从事什么工作，等等，显然是每一个从事考试理论研究和考试实践工作的人必须弄清楚的问题。而这些问题实质就是人才的定义、人才的标准和人才的分类等。所以，重视和加强人才标准、人才定义、人才分类的研究，决不仅仅是

人才学的事，也是考试学的重要任务之一。

一、人才的定义

关于人才的定义，目前尚处于讨论、研究阶段，界说多种多样。诸如：叶忠海等在《人才学概论》中指出："人才，是指那些在各种社会实践活动中，具有一定的专门知识、较高的技能和能力，能够以自己的创造性劳动，对认识、改造自然和社会，对人类进步作出了某种较大贡献的人。"①王康、王通讯等人，在《人才知识手册》中又将人才界定为："人才是人群中比较精华、先进的部分，是人民群众推动历史前进的代表。"②而《辞海》则把人才定义为："有才识学问的人；德才兼备的人。"③此外，见诸于各类书刊的定义还有："人才是成才的内在多因素的有机结合体"，人才是"对四化作出贡献的人"，"智能较高的人"，"知识结构超群出众的人"，等等。

以上种种关于人才的定义，虽然表达形式不同，强调各有侧重，但都从不同角度反映了人才内涵的基本要素，凡人才者，必须具备创造能力；必须具有比一般人更丰富的知识与才能；应对人类社会的进步和发展起积极的促进作用。

一个正确的概念，应该准确、严密和科学地反映某一事物（或现象）的本质特征。为了给"人才"这一概念下一个比较准确、比较科学的定义，我们先简略地讨论一下人才的本质特征。

现代人才学家研究的结果认为，人才的本质属性有以下三点：

其一，创造性。即人才较一般人具有更丰富的知识和较强的

---

①叶忠海等：《人才学概论》，湖南人民出版社1983年版，第59页。

②王康等：《人才知识手册》，湖北科学技术出版社1985年版，第1页。

③《辞海》，上海辞书出版社1979年版，第302页。

创造能力，能进行创造性劳动并能较一般人在某一方面或几个方面作出更大的贡献。

其二，进步性。即人才是人群中的优秀部分、先进部分，对社会发展和人类进步总是起着某种推动和促进作用。所以，即使是十分有才华的人，如果他是逆历史的潮流而动，也不属于人才之列。

其三，社会性。作为人群中比较优秀部分的人才，总是以一定的方式存在于社会之中，并受到一定社会关系的制约。所以，人才是一个历史范畴，无不打上时代的烙印。

根据人才所具有的本质属性或本质特征，我们将人才定义为：具有丰富知识和一定创造能力的，能代表所处时代发展趋势的，并能为社会的物质文明和精神文明作出贡献的人。

二、人才质量的基本标准

人才质量的基本标准包括思想品德、智能和体质三个方面。

（一）人才质量的思想品德标准

“德者，才之帅也”。所以，在人才质量的基本标准中，道德品质具有决定性的意义。这不仅是人才的进步性所要求的，而且也是人才定义本身所决定的。因为“才胜德者，谓之小人”，“小人挟才以为恶”，而“挟才以为恶者，恶亦无不至矣！”①人才质量的道德品质标准主要包括以下四个方面的内容：

1．政治品质。

这主要是指一个人的政治态度、政治立场。不同的时代，不同社会制度的国家，不同的阶级，对人才政治品质的具体要求是不一样的。如在我国现阶段，是否承认并坚持四项基本原则——坚持党的领导、坚持社会主义道路、坚持无产阶级专政、坚持马列

①司马光：《资治通鉴》。

主义毛泽东思想，便是人才政治品质的主要标准。

2．伦理道德。

这主要是指一个人在处理人际关系和其他社会关系方面的道德情操和行为准则。不同的阶级，有不同的伦理道德标准。在社会主义国家中，伦理道德的基本标准主要是：热爱祖国、热爱集体、尊师重友、团结同志、谦虚谨慎、心胸开阔、待人热情、诚实等。

3．心理品质。

所谓心理品质，系指包括兴趣、气质、性格、习惯、意识、意志力等品质的个性心理特征。有无强烈的事业心、顽强的拼搏精神、高度的责任感和探索未知的好奇心等，是检验人才心理品质的基本标准。

4．世界观（或哲学观）。

这主要是指一个人对人类社会、自然界乃至宇宙的基本看法。具体地说，就是用什么样的哲学观点去观察和解释人类社会，观察和解释自然界以及观察和解释宇宙的发生、发展和演变。不同的世界观（哲学观），对这一切的认识和解释是不同的。在社会主义国家中，提倡的是历史唯物主义和辩证唯物主义的世界观（哲学观）。

简言之，人才思想品德应该具备：为坚持真理的献身精神；为坚持进步的不妥协性；为坚持正义的斗争性；对祖国、对集体、对同志的满腔热忱以及能运用历史唯物主义和辩证唯物主义解释社会现象、自然现象。

应该指出的是，人才道德品质的四个方面，是互相关联、互相影响又彼此制约的一个整体。

人才思想品德属于意识和观念性的东西，对其进行甄别和检测，仅凭书面文字形式是不够的，应把书面文字考试与实际表现的考察有机结合，并以实际表现的考察为主。因为一个人的言论

和行动是受其观念和思想意识支配的，是观念和思想意识的外在表现。

（二）人才质量的智能标准

人才质量的智能标准是人才的创造性所要求的，是人才区别于一般人的根本标志。

人才的智能标准包括以下四个方面的内容：

1．获取知识的能力。

把获得知识的能力而不是把知识的占有量视为人才智能的重要标准之一。这是因为在现代社会中，一方面知识量急剧增长，另一方面部分知识老化的速率加快。在这种情况下，即使是一般人，也不能一劳永逸地占有知识，人才由于创造性劳动的要求，更是如此。所以，获得知识能力的大小应是衡量人才智能的重要标准之一。

2．创造的能力。

所谓创造的能力，主要是指一个人运用所获知识去进行创造性劳动的能力。它包括逻辑思维能力、分析综合能力、观察能力、实验设计能力等。一般说来，在知识占有量相同，且各种投入相等的情况下，创造性劳动的产品（新知识，新实物）数量多、质量高，对促进科学技术发展和社会进步的意义大，其创造能力也大。反之，则小。

3．表述的能力。

任何创造性劳动的结果，只有让人们了解、为社会承认才能达到目的，才富有实际意义。能否完整、准确而又生动地表述，是能否达到这一目的的基本条件。所以，表述的能力，也是人才智能标准的内容之一。

表述能力的体现形式主要有两种：口头表述和书面表述。既善于口头表述又善于书面表述的人，其表述能力最强。仅长于其中一种表述形式的人次之。既不善于口头表述又不善于书面表述

的人则其表述能力较差。

4．组织管理（包括交往）的能力。

除少数特殊的情况外，组织管理能力也应成为人才智能标准的内容之一。因为在具有高度文明的现代社会中，各行各业各种活动几乎都离不开管理。很难设想，在当今的社会中，一个完全脱离管理、毫无交往的人，尚能进行有效的创造性劳动。

以上关于人才质量的智能标准中，研究创造的能力是最主要的。

（三）人才质量的体质标准

除体育等专业人才在体质、体能方面应具有特殊的要求外，对于其他类型的人才，一般要求体魄康健，具有能适应本职工作的体质和体能。

以上关于人才质量的基本标准，是从人才个体的角度上提出来的。对于一个国家、一个地区、一个行业或一个单位来说，还有一个人才群体的质量标准问题。

关于人才群体的质量标准，除以上三条，即人才质量的道德品质标准、智能标准和体质标准外，还需加上反映人才群体结构方面的标准，这就是：人才群体结构应适应两个方面——所属系统（可以是一个国家、一个省、也可以是一个行业、一个单位）和人才群体自身。

人才群体与所属系统的相适度包括以下内容：（1）人才群体的数量与所属系统的需求比（比值为1时最佳、比值大于1或小于1时，则应进行调整）；（2）人才群体的专业分布与所属系统需求的相适情况；（3）人才群体的实力（主要是创造能力）与所属系统需求的相适情况。只有当这三个方面都满足所属系统的需求时，该人才群体的质量才是高的。

人才群体内部结构包括三方面的因素：（1.）年龄结构。合理的年龄结构应该是老、中、青的比例比较合理；（2）级别结

构。所谓级别，是指高级人才、中级人才和一般人才。合理的级别结构应该是高级、中级和一般人才保持一定的比例；（3）专业结构。专业结构合理与否，取决于专业结构是否符合所属系统专业分布的需求。

人才群体结构标准的两个方面是密切相关、互相影响的。一般说，凡结构合理的人才群体，其质量都是较好的，与人才群体质量标准的其他三条（道德品质、智能和体质）相比，群体的结构标准更能反映人才群体的质量，更具有综合性，因而也更重要。

三、人才的分类及原则

科学的人才分类原则，科学的人才分类，可以揭示人才群体的特征，也可以揭示不同类别人才的个性。而弄清人才群体的特征，尤其是弄清不同类别人才的个性，是提出考试理论、考试方法、考试种类和各种考试内容的基本依据之一。

（一）人才分类的一般原则

现代人才学理论认为，人才的分类方法可多种多样，但任何一种科学的人才分类方法，都应遵循以下五项原则。

1．客观性原则。

一事物区别于它事物的标志是其特殊性或个性。在人类社会这个大千世界中，人才济济、种类极繁。但从整体上讲，人才有其共同的属性和特征，这是人才作为群体区别于一般人的特殊性，也是人才群体的个性。然而，在人才个体，即组成这个群体的成员之间，显然各有千秋——知识结构上的不同，特长、能力上的差异等。例如，同属人才之列的将军、作家、物理学家、歌唱家、运动健将、特级厨师、医学家、建筑学家……他们在知识结构、特长、能力的表现形式等方面，都各不相同。正是人才群体内部各类成员之间知识结构、特长、能力表现形式等方面的特殊

性，使人才的分类不仅可能，而且必要。也正是这种特殊性，为人才的分类提供了最本质的依据。所以，人才分类最根本的原则，就是人才的社会存在的客观性。具体地说，人才的分类必须符合社会中人才群体的实际情况。社会中实际活动着哪些个性的人才，在人才的分类中就应该有反映那些个性的类别。同样，社会实际中不存在的，人才分类的类别中也不应该出现，一句话，人才的分类，一定要与社会中实际活动着的人才群体的实际情况一致。

2．整体性原则。

人才是人群中的精华部分，是一个整体，任何关于人才的分类，都必须是对这个整体而言，否则就毫无意义。

所谓人才的分类，是对人才这个整体内部的层次性而言的。人才整体是由社会各个领域、各行各业的优秀分子组成的。人类社会活动领域的广泛性、复杂性和多样性，决定了人才整体必然象一个“百花园”。百花园中的百花，虽然朵朵娇艳、千姿百态，但只要按一定的标准，总是可以将它们分成若干类的。如按颜色分，花有赤、橙、黄、绿、青、蓝、紫之别；按植物学的分类方法，花可分为木本、草本等；此外，还可以有其他的分类方法。人才这个整体的情形也正是这样，可以按照一定的标准，将其分成若干类别。例如，气质、性格、能力、知识结构、研究方向、社会承认先后等等，都可以作为对人才整体进行分类的标准。总之，人才整体是有序的，是可以分类的，并且任何关于人才的分类，都只能是对这个整体而言的。

3．相对性原则。

在组成人才整体的要素中，这样的情况是屡见不鲜的，即某一社会成员既是指挥千军万马的将军，又是在文坛上享有盛誉的诗人，甚至还可能是运动场上的健将或棋坛上的高手。在按人才整体分类时，他既可划在军事人才类，也可列入文学人才之列，

还可列为体育类人才。由此可见，人才的分类只能是相对的。在决定人才群体中某一个具体成员的属类时，只能是相对于我们所选定的标准，取其最突出的特征而定。所以，在人才的分类中，切忌绝对化。

4．动态对应原则。

动态对应原则包括两个方面的要求：一是人才整体中人才的类别应该是动态对应的；二是对具体人的类别归属，也应是动态对应的。

人才群体中有哪些类型的人才，任何时候都是由社会发展的状况所决定的。社会总是处于不断的变化、发展之中，总是不断地出现一些新的行业、新的部门、新的领域，同时淘汰一些旧的东西。伴随着这些新的行业、新的部门、新的领域产生，则又将涌现出一批批与之对应的新型人才。而与那些被淘汰了的东西相对应的人才类别，则随之成为历史。因此，人才永远是一个历史范畴，人才群体的结构总是随社会的发展而变化。无论那一种人才分类方法，都应该也只能是动态的。否则，就不能与社会人才群体的真实情况相对应。

在确定人才群体中具体成员的类别归属时，也要注意动态对应的原则，因为这些成员不可能终生毫无变化。比如，鲁迅早年学医，这时的鲁迅属人才群体中的医学类人才。后转学文，并成为中华民族的一代文化旗手，如仍按其最初的人才类别归属，显然是不行的。这种情况虽然并不具有普遍意义，但在人才群体中，以及人才史上也不乏其例。

5．有利于管理和考试的原则。

对人才群体进行分类的目的和意义，在于使人才群体中的每一种人才乃至每一个人才都能充分地发挥作用，也即人尽其才，才尽其用。而要做到这一点，对人才群体施以有效的管理是必不可少的。考试是人才管理的重要手段之一。任何人才群体的最初

分类固然离不开考试，更重要的是，在此之后的对各类人才的管理也离不开考试。因为在同一类别的人才中，在"能级"上仍然存在着差别，且各人才个体的"能级"也是动态的。要做到人尽其才，才尽其用，或换一句话说，使每个人才个体所担负的工作总是与其"能级"相一致，就必须定期或不定期地对他们进行考试或考察。所以，在进行人才分类时，必须注意有利于人才的管理和考试。

以上五项是人才分类的一般原则，特别是其中的客观性原则和整体性原则，对于人才的分类具有重要的指导意义。此外，在对人才进行分类时，还应考虑到如何有利于人才的开发和利用等因素，兹不一一赘述。

（二）人才的分类

现代人才分类的方法日趋多样化。其中较有代表性的是，法国施普兰格尔的人才分类法、日本的宫城音弥的人才分类法、按能力类型的人才分类法、按研究方向的人才分类法、按知识结构的人才分类法、按性格类型的人才分类法，等等。

纵观当今众多的人才分类方法，我们认为纵横方向综合的分类方法，是一种既符合前面提出的人才分类原则，又综合、借鉴了古今中外人才分类研究成果，并能比较客观地反映社会人才实际情况的人才分类法。因此，也是目前用得比较广泛的一种人才分类法。

这种方法的特点是：横向，以社会的专业分工作为人才的分类依据。社会分工是历史发展的产物。各种社会职业的出现，它反映了一定时代社会发展的状况，而且不同社会职业的存在，又必然有其独立存在而彼此不可互相替代的内在根据。众多的社会职业的存在，便使社会成为有机的和谐的整体，也正是这些社会职业的变化，推动了并正推动着人类社会的发展。所以，从职业分工的角度来研究人才分类，既可以通过揭示职业的特点来研究

各类人才的特殊性，使人才的研究更符合社会要求的实际，又可以从整个社会发展的角度，在更广阔的范围内把各种类别的人才有机地统一起来。纵向，或以人才的知识水平、创造能力为依据，分为高、中、初三个层次（每个层次还可以再分），或以人才所取成果的社会影响为依据，分为国际水平、全国水平、地区（系统）水平（每层次也可以再分）等。因此，这种纵横方向综合的分类方法，它具有客观性好、直观性强、可行性大的突出优点。

现在，我们试运用纵横方向综合的分类方法，简略地讨论一下人才的类别。考虑到横向的层次划分比较简单，故从略，仅就纵向的分类进行一些探讨。

从社会分工的情况看，现代人才可以粗略地划分为五大类：科技人才、军事人才、管理人才、文艺人才和体育人才。这五大类人才都可以作多层次的细分。这里我们只将科技人才作进一步阐述和细分。其他仅作简要介绍。

（1）科技人才。王康、王通讯在《人才知识手册》中对科技人才作了这样的界定："在社会科学技术劳动中，以自己较高的创造力，科学的探索精神，为科学技术发展和人类进步做出较大贡献的人。"①我们认为，如果这里科学技术是广义的，则这个定义是比较完整和准确的。

在现代，科学和技术已经是两个既密切相关而又不完全相同的概念和领域。科学是关于自然、社会和思维的知识体系。其目的是解释自然、社会和思维现象，揭示自然、社会和思维规律。技术则是运用科学理论或经验，按照人们的要求创造出改造客观世界的方法和手段。其目的在于对客观世界的控制和利用。所以，对科技人才分类的第一层次是，科学人才和技术人才。

①王康等：《人才知识手册》，湖北科学技术出版社1985年版，第151页。

科学是数学、自然科学、社会科学和思维科学的总称，故科学人才又可以分为数学人才、自然科学人才、社会科学人才和思维科学人才。

数学、自然科学、社会科学和思维科学，仍然是包罗万象的学科群。如自然科学，即使按传统的分类，被人们称为一级学科的就有物理学、化学、天文学、生物学和地理学。所以自然科学人才又可分成物理学人才、化学人才、天文学人才、生物学人才和地理学人才。这些一级学科实质上都是一个个庞大的学科家族，它们都各自拥有几十个乃至更多的二级、三级学科。而一个学科就有一种对应的人才类别，所以，从纵向看，根据不同的学科和层次，科学人才可以细分为与各层次学科数相对应的类别。图4－1是科学人才按纵向不同层次分类的示意图。通过图4－1我们可以直观看到科学人才与学科（从社会分工的角度讲，则是科学人才与社会专业分工）一一对应的情况。

技术人才纵向分类与科学人才纵向分类的情形基本相同。

（2）军事人才。军事人才是在军事领域中能比一般军人作出更大贡献，品质高尚，并且有正义感的人。军事人才的一般特征是：在军事领域中具有一定的专长和一定的创造能力，并为一般的军人难以期及；爱憎分明，一切都是为了人民的事业和正义的战争。军事人才是一种人数众多，种类纷繁，于国家于民族具有特殊意义和特殊作用的人才，也是古今中外国家统治者至为关注和重视的人才。

按照不同的分类座标，军事人才的分类也不同。常用的分类方法有三种。按部门分：军事指挥人才，军事参谋人才，军事政治人才，军事后勤人才等；按兵种分：陆军军事人才，空军军事人才，海军军事人才，炮兵军事人才，通信兵军事人才，侦察兵军事人才，装甲兵军事人才等；按能级分：高级军事人才，中级军事人才，一般军事人才等。

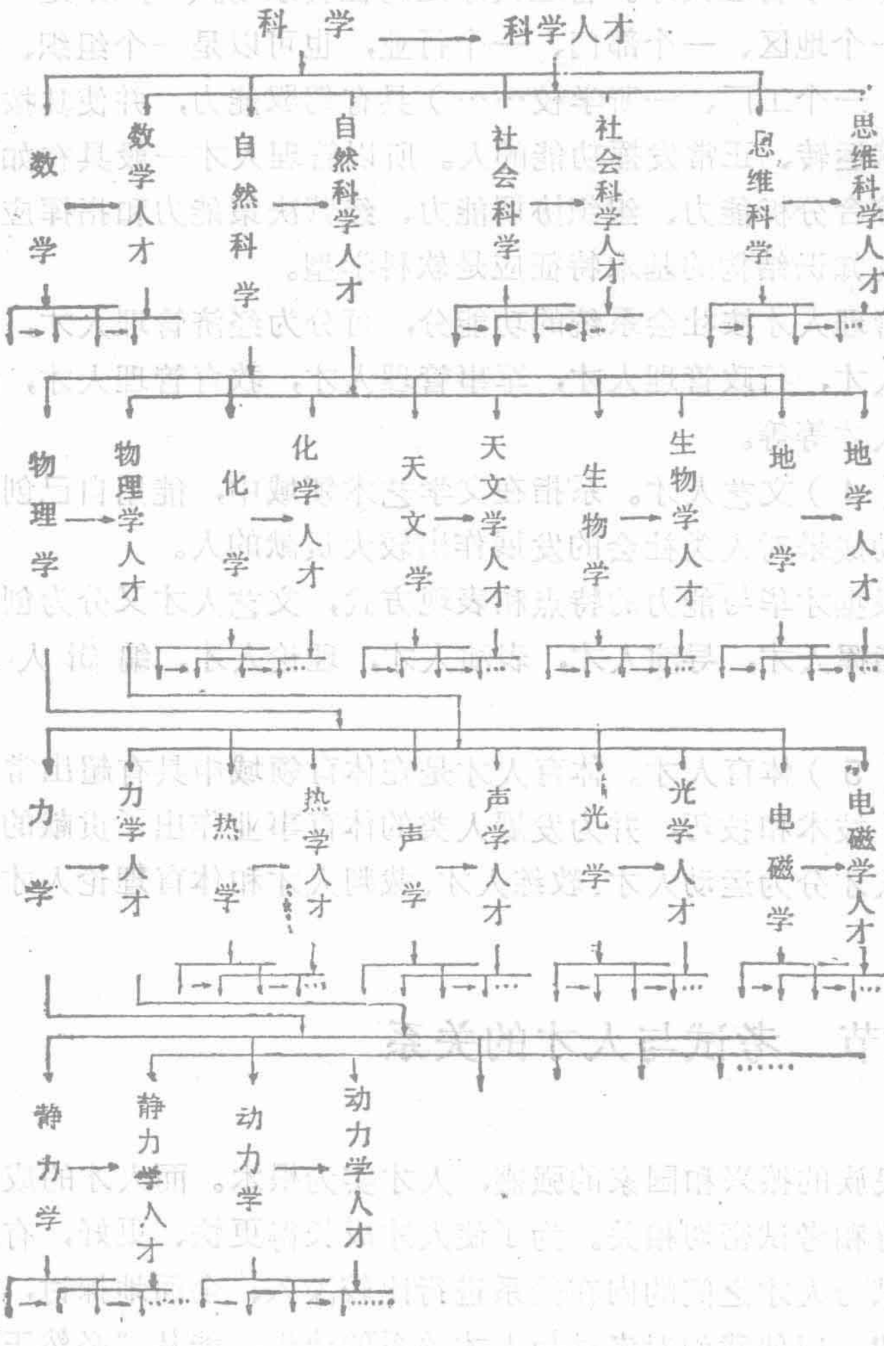


图4－1 科学人才纵向层次分类示意图

（3）管理人才。管理人才是对社会系统（可以是一个国家、一个地区、一个部门、一个行业，也可以是一个组织、一个团体、一个工厂、一所学校……）具有驾驭能力，并使其按一定的要求运转、正常发挥功能的人。所以管理人才一般具有如下能力：综合分析能力、组织协调能力、经营决策能力和指挥应变能力。其知识结构的基本特征应是软科学型。

管理人才按社会系统的功能分，可分为经济管理人才，科技管理人才，行政管理人才，军事管理人才，教育管理人才，企业管理人才等等。

（4）文艺人才。系指在文学艺术领域中，能用自己创造性的劳动成果对人类社会的发展作出较大贡献的人。

根据才华与能力的特点和表现方式，文艺人才又分为创作人才，指挥人才，导演人才，表演人才，理论人才，编辑人才等等。

（5）体育人才。体育人才是在体育领域中具有超出常人的技能、技术和技巧，并为发展人类的体育事业作出了贡献的人。体育人才分为运动人才、教练人才、裁判人才和体育理论人才等。

第三节 考试与人才的关系

民族的振兴和国家的强盛，人才实为根本。而人才的成长，与教育和考试密切相关。为了使人才成长得更快、更好，有必要对考试与人才之间的内在关系进行比较深入、全面地探讨，弄清其规律，以便我们对考试与人才关系的认识，能从“必然王国”逐步走向“自由王国”，并在考试实践中正确处理这些关系，以促进人才的健康成长。

一、考试门类设置与社会人才群体结构

这里所说的考试门类，是就人才的种类而言，意即考试的门类与人才的种类相对应。

从客观上讲，社会人才群体的结构，是由社会发展、社会分工的情况决定的。这是形成社会人才群体结构的一条基本规律。然而，在社会发展的某一个方面、某一个局部或某一个阶段，人才群体的结构又是可以控制的。例如，一个国家可以（而且应该）根据本国某一时期或某种特殊需要，重点培养一种或某几种人才而限制培养另一种人才。一个地区也是如此。但无论是一个国家还是一个地区，要对社会人才群体结构实行控制，最终都离不了考试这一手段。这就是：需要哪种人才，就设置甄选、录用哪一种人才的考试；重点需要哪一种人才，就对甄选、录用那一种人才的考试采取特殊措施（当然还有其他措施配合）；限制某一种人才，也对甄选、录用这种人才的考试采取特殊的措施；如果在某一段时间内有意不让某一类人才数量增加，还可以在相应的时间内停止甄选、录用相应类别人才的考试，等等。所以，甄选和录用人才的考试门类的设置，对一个国家、一个地区的人才群体的结构，具有相对的控制作用。

考试之所以对社会人才群体结构具有这种相对的控制作用，是因为考试对社会人才群体结构的形成具有导向功能和调整功能。

所谓考试对社会人才群体结构形成的导向功能，是说考试（再加上其他相应的措施）可以引导或影响正在成长中的人才对未来职业的选择和确定。当然，正在成长中的人才，他们对各自未来职业所作出的选择和确定，最终决定于社会的人才需求。考试对社会人才群体结构的形成具有导向作用，是由于以下两种原因：一是考试对重点选拔录用什么人才，限制录用什么人才，不招录什么人才，总是具有明显的倾向性，直接或间接地反映了社

会对人才的需求，这对于每一个有可能成才的人来说，无疑是一种信息、暗示或启迪。二是凡有可能成才者，总是希望其才能有用武之地，因而在其选择和确定自己成才方向或目标的时候，不能不考虑社会的实际。而在考虑社会实际时，由考试所反映出来的社会对人才种类的需求情况，便自然而然地产生作用。

考试对社会人才群体结构的调节功能，是借助考试对社会人才群体结构的导向功能实现的。考试对社会人才群体结构的导向功能与调节功能之间的这种关系是显而易见的，当社会当局认为社会人才群体的结构不合理时，就需要通过考试的导向功能来改变这种不合理的结构，使社会人才群体结构趋于合理。而这种由不合理的群体结构到合理的群体结构的变化过程，就是考试对人才群体结构的调整过程。这种调整作用，也就是考试对社会人才群体结构的调整功能。

由于考试门类设置与社会人才群体结构之间具有这样的关系，所以在考虑和决定考试门类的设置时，必须注意以下两个问题：

（一）社会人才群体结构的现状与社会发展实际需要的相适度

一般说来，如果社会人才群体结构的现状与社会发展实际需要的相适度高，或者说，社会人才群体结构的现状能较好地满足社会发展的实际需要，便说明当时的社会人才群体结构合理。从考试方面讲，则意味着（到当时为止）人才考试的门类设置及其措施是合理的。在这种情况下，现行人才考试的门类及其措施可以继续实行。如果社会人才群体结构的现状与社会发展的实际需要的相适度不高，甚至很低，亦即社会人才群体结构及其分布情况满足不了社会发展的实际需要，则说明此时社会人才群体结构及其分布是不合理的。从考试角度讲，则意味着现行人才考试的门类设置及其措施已不合理，至少是某些门类设置不合理。处于

此种情况，就应调整原来的考试门类及其措施，使之引导人才群体结构向适应社会发展需要的方向转变，直至能较好地满足社会发展需要为止。

（二）社会近、中期发展对人才群体结构的要求

弄清社会近期（10年左右）、中期（15～20年左右）发展对人才群体结构的要求，对于调整和确定人才考试的门类设置，保证培养和成长中人才群体的结构能适应十年、二十年后社会发展的需要，是十分必要的。因为无论是一种人才的成长，还是一种人才群体结构的形成，都需要一个相应的时期，一般在10年左右。所以在调整或确定人才考试的门类设置时，应考虑到十年、二十年间社会发展对人才群体结构的需要。

为了比较准确地弄清未来（10年或20年）社会发展对人才群体结构的需要情况，必须重视抓好社会发展情况的预测。由于预测的结果，是调整和确定人才考试门类设置的依据，故要求这种预测工作必须具有较高的科学性和准确性。

二、考试科目与人才个体知识结构

这里所讲的考试科目，是指与人才种类相对应的考试门类确定之后，对每一种人才进行考试的具体科目。例如无线电工程师是人才群体中的一种，为了区别于其他种类的人才，判定某人是否具备无线电工程师的素质与智能，社会人事部门往往是通过规定相应的知识结构，作为区别一种人才与其他人才的标准。至于某人是否属于这种人才，则需通过以该种人才知识结构为内容的考试来判别的。所以，考试科目的确定直接决定了人才个体应具备的知识结构。

为了使人才的知识结构能较好地适应社会发展的需要，在确定考试科目时应注意两点：

（一）当代科学技术的特征

人才与科学技术是分不开的。任何人才，都具有学科（或技术）的属性。也就是说，任何人才按其知识结构的最主要的特征，都可划归到相应的学科或技术领域。因此，充分认识当代科学技术的特征，并以此为确定考试科目的重要理论依据，对于提高人才个体知识结构与科学技术发展需要的相适度，是十分有益的。

当代科学技术最主要的特征有三：第一，既高度分化又高度综合。分化是形式、是现象，综合是本质。所谓分化，是指科学技术的学科越来越多，而且新学科尚如雨后春笋般地涌现。所谓综合，是指每一个新的学科的形成，都是几个以至几十个原有学科互相渗透的结果。科学技术体系中出现一个新的学科或技术领域，人类社会的分工中就必然增加一个与之相对应的专业；第二，科学一体化的进展非常迅速。本世纪乃至在第二次世界大战以前，社会科学、自然科学之间的"鸿沟"曾被认为是不可逾越的，即使在社会科学、自然科学内部的各学科之间也是界线分明，壁垒森严。但在第二次世界大战40余年后的今天，现代科学技术发展的滚滚洪流，既冲塌了社会科学、自然科学内部各学科间的壁垒，也使社会科学、自然科学之间的鸿沟处于消逝之中。即使从前被人们认为是"风马牛不相及"的学科（如语言学和数学），现在也结为亲家；第三，部分知识老化的速度日益加快。

（二）社会发展对人才个体知识结构的要求

社会在其不同的发展阶段，对人才个体知识结构的要求是不同的。当社会的文明程度不高，发展变化速度缓慢时，它对人才个体知识结构的要求的特点是一维、稳定和知识型的。而当社会的文明程度高，发展变化剧烈、迅速的时候，它对人才个体知识结构的要求则是多维、动态和能力型的。现代社会是一个急剧变革的时代，它对人才个体知识结构的要求由过去的一维、稳定和知识型转变为多维、动态和能力型，要求人才具有自学能力、研究能力、思维能力、表达能力和组织管理的能力。

总之，为了使人才个体的知识结构能较好地适应所处时代发展的要求，对人才个体知识结构的形式具有强烈影响的考试科目的选择和确定，必须注重人才所处时代的社会特点和实际需要。

三、考试内容与人才创造思维能力

创造思维能力，是人才区别于一般人的本质所在。因而创造思维能力的大小、强弱，则是衡量或评定人才质量的主要指标。而人才创造思维能力的培养和造就与考试有密切关系。

考试内容就其特点可以分为两大类型。一种是单纯知识型的。另一种则是综合启迪型的。单纯知识型的考试内容，属于静态的或“死”的知识。它偏重人的“记忆力”的测试。与之相适应的学习方法，便是死记硬背，与之相适应的教学方法则是“填鸭式”。这种考试只能培养和锻炼人才的记忆。综合启迪型的考试内容，则是动态的或“活”的知识，如提出问题，要求找出解决的方法；摆出现象，要求解释其前因后果，等等。这种类型的考试内容不仅对人们的记忆力有一定的培养和训练，更重要的是着眼于对人才的创造思维能力的培养和训练。

由此可见，考试内容的着眼点不同，对人才的培养、训练的结果也就不一样。如果是甄别人才，单纯知识型的考试内容，它所测度、甄别的只是人才的记忆力和静态的知识，而综合启迪型的考试内容，它所测试的不单是人才掌握知识的数量，更重要的是检测人才运用所学知识去解决实际问题的能力。正确认识考试内容与人才创造思维能力之间的这种关系，有利于培养人才的创造思维能力，也有利于对人才创造思维能力的判别。

第五章

考试与考生

考试是主试与被试两种主体的交互活动。任何考试的设计与实施，只有严格遵循人的身心发展规律，才能获得成功。

考试的对象是人，而人的素质和智能是其生理和心理现象的外化，与自身生理和心理的发展密切关联。因此，考试必须遵循人的生理和心理发展的规律，才能达到预期的效果。

人，既是生物实体，又是社会存在物。人的生理和心理的发展，是在内因——遗传和外因——环境互相作用下逐步进行的。遗传素质是生理和心理发展的物质前提，没有这个前提，人的生理心理不可能得到发展。遗传素质是在环境影响下逐步发展成熟的。它的成熟过程制约人的发展水平和阶段性，外化为不同年龄阶段不同的生理和心理特征，它就是考试所需要依据的生理和心理发展的规律。遗传素质本身的个别特点以及环境与教育条件的影响，人在生理和心理上形成了鲜明的个别差异。这就为考试的甄别和测度提供了可测可比的客体，使考试的存在和发展有了坚实的客观基础。

遗传素质在人的生理和心理发展上起着重要的作用，但它不能决定人的发展，仅仅是为人的发展提供了物质前提和可能性。只有通过后天的环境和教育的影响作用，以及人本身的主观能动

的努力，使遗传素质得以发展，才能成为人的现实属性，形成知识、能力、思想、品德以及个性和健康体质等等。考试成绩的优劣，是人在后天所获得的知识能力和形成的思想品德及个性外化的表现，而不是遗传素质的优劣。

所谓环境，即是人们生活的客观世界，包括自然环境和社会环境。对人的生理和心理发展起决定作用的，是人类所特有的社会环境，其中教育起着主导作用。社会环境包括家庭、邻里、亲戚、朋友、学校以及娱乐、劳动场所和风俗习惯等等。环境对人的影响具有二重性，既有积极的教育因素，也有自发的消极影响，对人的发展可导向正确的道路，也可将人的发展引向反面。只有环境在教育的作用下，抵制自发的消极因素，引入社会所需要的方向，并接受积极因素的影响，对人的发展才具有积极的意义。因此，教育（学校教育或社会组织的教育）对人的发展起着主导作用。考试是根据社会的要求，对人的素质与智能进行检测，它与教育在目的上具有一致性，考试不能脱离教育和社会环境来甄别与测度人在德、学、才、识、体诸方面存在的个别差异。也就是说，考试要根据教育或社会环境提供“产品”的实际状况进行检测，否则就会失去考试的意义。

对于人的发展来说，遗传素质、环境和教育的作用是辩证统一的，并由此形成了人的自身生理和心理的发展规律，在不同的年龄段表现出一些不同的特征。考试必须遵循人的发展的普遍规律进行测试，才可能反映考生的素质和智能的真实状况。当然，客观事物除存在一般矛盾或一般规律外，也存在特殊矛盾或特殊规律。人的发展也是如此，大多数是普通智能者，但也存在超智者和特殊才能者以及身残者，他们与一般人在身心上有更显著的差异。因此，考试又必须根据具体的对象，按照特定的原则，采用相应的方法进行甄别和测度。探讨考试和考生的关系，就是要从考试的角度，依据考生生理和心理发展的规律，探讨考试的原

则和方法；阐述考生家庭、社会环境与考试的关系及其影响。

第一节　考生的身心发展与考试

从出生到成人，人的生理和心理都经历着一个发展过程，在发展过程中，不同的年龄阶段有相对稳定的不同特点，即年龄特征。它是人本身的内因和环境的外因相互作用发展形成的。人的发展既有连续性，又有阶段性，年龄特征既是稳定的，也是可变的，人与人之间还存在着个别差异。因此，进行考试既要遵循人的生理和心理一般规律，又要考虑人的个别差异，并采取相应的对策，才能达到预期的目标。本节将分别阐述儿童考试、青少年考试及成人考试的特点、原则与方法。

一、儿 童 考 试

儿童考试系指童年期小学儿童的考试。

（一）小学儿童身心发展的特点

小学儿童身心正处于发展之中，生理机能和心理功能均不能承受过重负担，是儿童考试必须注意之点。因此，特将儿童身心发展状况叙述于后，以作为儿童考试的依据。

儿童肌肉处于不断增长之中，收缩力量小，伸展性和弹性也差。因此，肌肉工作的耐力差，容易疲劳，但儿童新陈代谢旺盛，肌肉工作疲劳后，恢复也较快。

儿童血液总量比成人少，血液中红细胞和红蛋白量也比成人少。因此，血液的输氧功能比成人差，儿童活动时间不能过长，活动量不能过大。

儿童的心脏功能和神经系统对血液循环系统的调节不及成人生活，生理性心律不齐”的现象也较多。儿童学习和体育活动及

劳动的时间和负荷量均不能超过其生理承受量，否则会导致儿童生理性心律不齐转化为病理性心律不齐。

儿童的神经系统的发育与其他器官、系统相比，在结构和机能上是较早的，但大脑功能尚未完善，兴奋与抑制的平衡性较差，转换较快，灵活性也好，逻辑思维能力开始发展。因此，能学习较复杂的科学知识和技术。但儿童好胜心强，往往过高估计自己。所以，施考应特别注意他们负荷量的控制，防止过度疲劳而造成神经细胞的损伤。

儿童感知事物比较笼统，只注意到一些孤立的现象，看不出事物之间的联系，时间和空间观念比较模糊。无意注意占优势，对直观的具体事物或新奇的刺激物，容易引起注意，但注意力不稳定、不持久，容易转换，缺乏自觉性。

儿童在小学低中年级时，形象记忆、机械记忆、无意记忆占优势。到了小学高年级，形象记忆和抽象概念记忆的差别逐渐缩小，逐步学会理解记忆，能提出记忆的目标，努力识记有关的知识。

小学低年级的儿童，想象以幻想为主，常把幻想当作现实。到三年级以后，想象逐步过渡到较为精确、完整、符合现实为主的想象。

儿童的思维以具体形象思维为主要形式，逐步向逻辑思维过渡，并开始出现初级的创造思维。

儿童情感处于较为平静、持久的愉快状态，社会化的高级情感发展很快，表现为强烈的集体荣誉感。意志品质的自觉性逐步提高，但自制力还不成熟，易违犯纪律。兴趣广泛而缺乏专一性。

总之，儿童的身心处于发展之中，考试应遵循其规律，使考试的形式、内容、方法、时限等符合儿童身心发展的实际。

（二）儿童考试的原则与方法

儿童考试的原则与方法，是根据儿童生理和心理发展的特点阐述儿童考试问题，重在儿童考试的特殊性，对于考试的共性问题不赘述。

1．儿童考试的基本原则。

小学儿童的考试，从教育行政的角度而言，是督导教育方针的贯彻落实，检查教学质量；从教师的角度讲，是检查学生学业成绩，评估教学效果，诊断教学中的问题，进行教学改革；从学生的角度说，是检验学习成效是否达到国家规定的要求。因此，小学儿童考试当遵循下列基本原则：

（1）针对性原则：考试必须依据国家的教育方针，按各科教学大纲的要求命题，符合小学儿童生理和心理发展规律。各个年级的考试、同年级不同类型的考试，都应各有区别。考试标准要相对稳定。

（2）激励性原则：考试是督导学校贯彻落实教育方针的手段，目的是提高教师的教学质量和学生的学习成绩。考试应遵循激励性原则，调动师生教与学的积极性，催其奋进，才有利于教育方针的贯彻；若用考试对师生进行卡和压，将会产生严重的消极作用。

2．儿童考试方法的特点。

小学是基础教育，目的在于提高民族素质，为进一步学习做准备。因此，小学儿童的考试，主要是合格考试与诊断考试。在考试的指导思想、组织与实施方面，务必讲究方法。

（1）思想与态度。考试是为了督导教育方针的全面贯彻，提高教学质量。因此，师生应本着实事求是的精神，积极、严肃、认真的态度参加考试。要摒弃考试前加班加点、突击复习、猜题、押题、圈题、划范围的不良作风；杜绝考试中暗示、提示、甚至舞弊的枉法行为。也要纠正按例行事，不重视，不准备，毫无进取心的消极心理状态。应遵循激励性原则，鼓励师生积极进取，按

照教学计划的规定，认真组织考前复习，从教与学的角度全面总结，争取圆满实现教学目标。

（2）组织与准备。命题是考试的关键。试题既要反映大纲要求、切合教材内容实际，又要符合儿童生理和心理的特点，具有针对性和强烈的情绪色彩；既有检验机械识记和理解识记的试题，又有再造想象和创造想象的试题；既有具体形象思维和抽象逻辑思维的试题，又有一定量的创造思维的试题。根据年级的不同，低年级应以机械识记、理解识记、再造想象、具体形象思维试题为主；中高年级则以理解识记、创造想象、抽象逻辑思维和一定量的创造思维试题为主。试题的匹配与各种类型试题的比例，要考虑考试的目的与教材的实际内容，视具体情况决定，既不拔高又不降低标准。儿童考试切忌盲目搬用赞科夫的高难度教学原则，不要脱离儿童实际或超过儿童承受力的极限。

小学儿童的考试在次数与时间上应严格控制。现在考试频繁，所谓单元考、摸底考、期中考、期末考，以及来自上级的统考、检查考试等等，名目繁多，而且缺乏严密组织，使师生忙于应付，考试质量不高，在某种程度上失去了考试的意义。儿童考试不仅应该控制考试的次数，而且要求在考试前有充分细致的准备，一自测的单元考试每学期有1～2次即可，期中、期末考试按国家规定进行，摸底之类考试应该取消。

目前，小学儿童考试的测试时间过长，少则40分钟，多则90分钟至120分钟。这不符合儿童生理和心理特征，有损他们的健康。一般说来，低年级每次考试的时间，应控制在30分钟以内，中高年级的考试，每次以不超过50分钟为宜。

儿童考试试场的设置应符合儿童生理与心理的要求。桌椅应规范化，桌子的高度一般以儿童坐在椅上整个身高的2／3为宜，坐椅则以儿童的两脚自然踏地为标准，桌面必须平面无光。试场墙面与天花板应为白色，不应张贴标语与宣传画。场内采光适

度，空气流通、冬季门窗关闭后，应留有气窗。试场周围环境要保持安静，不得有闲散人员走动或播放音乐。

考试试卷的准备是考试的重要环节。目前除重视命题外，试卷的制作却极为粗糙，如纸质差，字迹不清，漏字错题等现象常有发生。儿童考试的制卷应严格把关，各科试卷当采用不同颜色、质量较好的纸张，文字与图形除大小适度外，应字迹清晰、图形规范，试卷清样必须反复审查校对，严防错漏。

（3）施测与评讲。施测是考试的中心环节，成败在此一举。对小学儿童而言，测试好坏关键在于临场发挥。若怯场或分心，都影响儿童的测试成绩。而儿童在测试过程中能否正确发挥其聪明才智，主考人员往往起着重要的作用，如果主考者态度过分严肃，将给儿童应试心理以压抑感，造成儿童怯场；如果主考者态度随便，穿戴异常，又会分散儿童应试的注意力，从而影响儿童水平的正常发挥。因此，儿童考试的主考人在态度上应既严肃又亲切，其仪表需端庄而朴实。语言要简炼明确而亲切。还需注意的是，试卷评讲应从激励性原则出发，不能抱有恨铁不成钢的心理，对失误的考生或差生进行批评指责，对成绩优秀的学生也不能一味鼓励和肯定成绩，而不如实指出存在的问题与努力的方向。对于儿童测试成绩的评讲，教师应高姿态自咎，对失误者或差生寻找成功之处，挖掘闪光之点，鼓起他们前进的信心和争取优良成绩的勇气，同时帮助分析失败或差的原因，帮助他们制订学习措施，并加强督促检查。对考试成绩优秀者理应给予鼓励，但更重要的是帮助他们分析成功的原因，提出更高的要求，激励他们更上一层楼。对教师来说，自咎是进行反思，不断改进教学内容和教学方法，以便对儿童进行启发诱导，指引他们进行自我探索，掌握知识、发展能力，促使他们全面发展。

总之，小学儿童考试必须立足于基础教育，着眼于提高民族素质，目的在于全面贯彻教育方针，提高教学质量，为儿童继续

学习奠定基础，以保证他们在德育、智育、体育、美育及劳动技术教育诸方面得到健康发展。

二、青少年考试

青少年期包括少年和青年两个阶段（11、12～23、24岁），是儿童发展为成人的过渡时期。在这个时期，既长身体又长知识，世界观在形成，道德品质在定型；生活环境由家庭扩大到社会，对家长的依附性降低，生活逐步走向独立，是人生的非常时期。正确认识和掌握青少年的身心发展规律，明确青少年考试的原则和方法，对社会培养、选拔人才具有积极的现实意义。

（一）青少年身心发展的特点

青少年期的前期恰在生理发育的第二次高峰（第一次高峰是从成胎到两周岁），他们的身高体重、胸围、肌肉、骨骼大体接近成人标准。脑的结构和形态基本发育成熟，大脑皮质机能的发展也达到了相当的水平，两种信号系统相互协调，第二信号系统已起着主导作用。因此，青少年有了承受认识和解决复杂问题的能力，处于记忆的全盛时期。抽象思维从经验型向理论型转化，思维活动的独立性和批判性明显增强，但有偏激和片面的毛病。青少年的想象富有创造的特色，并由幻想多于理想逐步发展为以理想为主导。情感的自我调节能力不断增强，世界观处于形成时期，社会化道德情感已得到稳定的发展，对理想的追求更加自觉和主动，开始自觉地考虑未来的职业，愿为理想去拼博，乐于独立地去完成任务，并能善始善终。

基于上因，青少年个性倾向从不成熟逐步发展成熟，开始有较强的自我意识和独立评价能力，个性特征的心理系统渐渐形成。他们总是以满足和实现新的愿望为目的，对环境、教育和考试等作出相应的选择。

（二）青少年考试的原则与方法

从青少年生理和生理发展特点看，他们的体质和心理状况已完全可以接受或承受国家、社会团体和学校的考试，同时青少年生理和心理发展规律也为科学设计和实施考试提供了客观依据。

从生理角度看，青少年的心脏机能能保证血液循环系统正常运转，提供大脑所需要的营养和氧。骨骼、肌肉发育，特别是小肌肉群的发育，有力量进行较长时间的手的细致书写、演示、表演以及消耗体力的耐力、速度、技巧等测试活动。呼吸系统也能适应紧张的脑力劳动和体力活动，其肺活量已接近或达到成人的水平。在神经系统方面，大脑形态和结构发育已经完成，两种信号系统能相互协调，而且第二信号系统已起着主导作用。大脑能进行紧张的思维活动，利用储存的信息，解决高难度的认知问题。因此，在生理上，青少年已有承受各种考试负荷的条件。

从心理角度看，知觉能有目的有系统地观察物事1小时35分钟至3小时，为考试时间提供了保证。青少年是记忆的全盛时期，能适应知识的广度和深度的测试。青少年抽象逻辑思维由经验型向理论型的转化，以及聚合思维和发散思维的发展，便为承受多种内容的考试奠定了基础。再者，由于青少年道德感、理智感已趋成熟，自觉性、自制力、坚韧性等品质稳步发展，这有利于强化青少年对考试的责任感，使他们有毅力完成考试的测试内容。

从青少年生理和心理发展的特征与考试的关联分析，对青少年进行考试完全是可行的，青少年身心的发展规律不仅为考试提供了可遵循的科学依据，而且青少年素质与智能的个别差异，是可以通过考试进行甄别和测度的。

为青少年生理、心理的发展特点所制约，青少年的考试也有其特殊的原则和方法。

1．青少年考试的原则。

（1）　整体性原则。对青少年进行德、智、体、美、劳全面考核，知识和能力测试并重，这是青少年考试整体性原则的基本含意。

历史上任何国家与民族所进行的青少年考试，首先考虑的总是政治标准——德，其次是才。当今时代，青少年一代的考试，不管那种社会制度的国家，也是遵循这个原则的。但是，在人类社会的不同历史阶段，以及在同历史阶段不同社会制度的国家或民族，对青少年德、才的检测标准和内容是有别的。比如我国现阶段，随着生产和科学技术的发展，社会不仅要求青少年一代要有坚定的政治信念、良好的品德、强烈的事业心、宽厚坚实的理论知识和实践能力，而且要富有创造精神、和谐的审美情操、科学的管理素质和灵活的应变能力，以及健美的体魄。因此，考试也必须按照社会的要求，对青少一代的知识和能力进行全面测试。而青少年正处于发展之中，可塑性很大，只要有科学的教育和考试给予引导，他们是可以按照社会要求成长的。

（2）　激励性原则。青少年一代的考试，应对他们起鼓励鞭策作用，激励他们积极进取，促进他们全面发展，为实现崇高的理想而奋斗，而不能使他们对考试产生恐惧心理，束缚或压抑他们在德、智、体、美、劳几个方面的发展。因此，除对青少年全面考核外，必须坚持青少年一代在考试面前人人平等，鼓励他们进行正确思想指导下的合理竞争，以有利于促进青少年一代好学上进、健康成长为原则。同时，人的个别差异客观存在，青少年的素质和智能没有完全相同者。这种客观存在的个别差异，可以通过考试进行甄别，使他们各得其所，各尽其才，才尽其用。因此，通过考试鼓励青少年展开竞争，并在竞争中获得进步与提高，这对青少年个性的发展，社会人力资源的开发和利用，乃至推动社会前进，都将有益无害。

（3）　促进教学改革的原则。任何国家的人才都是通过教

育进行培养，培养的人才是否符合国家的要求，又是通过实践和考试来检验。由于教育具有滞后特点，培养的人才，可能在数量上与社会需求不一致，或在质量上与社会要求有一定的差距。因此，国家要通过行政手段或考试对教育进行调整，督导教育达到国家要求的标准。学校通过考试检查教学质量，也进行自我调整，以采用新的措施促进改革。教师通过考试诊断青少年学习情况，以便改进教学。可见，青少年考试不管是社会的，还是学校的，都与教育密切相关。要使考试给教学改革以积极的促进作用，就绝对不能脱离教育这个基础，必须把有利于促进教学改革和提高教育质量，作为实施青少年考试的基本原则。

2．青少年考试的方法。

青少年考试的方式与方法必须慎重选择，既要有利于青少年身心的全面发展，又要有利于知识、技能、能力、品质、体能等方面的准确测度和甄别，能客观反映每个青少年的真实水平。各种考试的方法，应与目的相符，不可随意施考，也不能套用某种方法，必须真正做到根据不同的考试目的选择有针对性的方法。譬如，下面列举的四类考试，因其目的不同，方法也就各不一样。

青少年的选拔考试，名额有限，竞争性强。如高校、高中、中专、中等技术学校的入学考试，以及工矿企业、事业单位的招工考试，一般采用团体考试与个别考试的方式。其方法可采用口试、笔试、实际操作及演示相结合的方法。值得注意的是，青少年考试临场经验少，怯场的情况难免发生。因此，不能完全依据一次考试成绩决定考生命运，必须全面了解考生的实际，并参考考生过去的成绩。对失误者应予再试的机会，以免埋没人才，同时对青少年也是一种激励。

青少年成就考试，目的是让青少年充分表现其才能，激励青少年进取，促使他们脱颖而出。因此，可以评价主题设计、专题论文、调查报告，以及创作、创造发明的成果等，与答辩相结合

的方式进行。亦可采用知识或技术演示竞赛的方法。这种考试应鼓励青少年自愿报名参加，不受名额限制，但必须有初试、复试、终试几个环节，使青少年的学业或技术成就能获得正确的评价，并给他们导以正确的方向，为其拓宽成才之路。

青少年合格考试，目的是检测青少年的学习和工作是否达到规定的标准。如初中、高中、大学的毕业考试，青年工人晋级考试，机动车辆司机驾驶执照的考试等。合格考试标准既定，针对性强，级别和专业有别。所以，考试方法应各个有异。学校系统的毕业考试知识和能力并重，但对不同级别的考生，方法也要有所区别。中学毕业考试一般采用笔试，人数少的还应进行口试。大学生的毕业考试，除上述方式方法外，可采用毕业设计或论文答辩的形式。青工的晋级或合格证书考试，基础理论知识的检测宜用笔试或口试，实际技能应采用操演测试法，或通过排除障碍、解决难题等方法进行实践性检验。

青少年的诊断考试，意在诊断青少年学习或工作的进展情况及存在的问题。其方法有调查、询问、学习成绩考查、工作质量检测等。有口头或书面的测试，也有实际操演或产品质量、作业的检查。

青少年考试的方式方法因考试目的和内容而异。所以，它必须根据青少年的特点，按照青少年考试的原则进行组织，考试的目的要明确，考务工作要严密，方法要得当，否则就不能测出被试者的真实成绩，也无益于青少年的发展。

三、成人考试

在考试发展史上，影响最大的是成人考试，如中国的科举考试和英国的文官考试，均属成人考试的范畴。成人考试在历史上是考试的主体，对儿童和青少年考试有程度不同的支配作用。

进入现代社会后，由于生产和科学技术迅猛发展，科学管理的实施以及人们对精神文明的追求，继续教育和终身教育的兴

起，成人考试更是发展空前，已渗透到社会各系统各部门。其门类也迅速增加，如教育系统的成人自学考试、各级成人教育的招生、毕业及平时学业成绩考试；社会系统各行各业的招聘考试、技术业务考试、晋级考试、职务聘任考试等。

成人考试随着年龄的变化和职业不同而各有其特点、原则与方法。

（一）成人各年龄段的特点

1．成年早期。（25～40岁）

成人的成年早期其生理和心理发育发展完全成熟，是生理机能和心理功能鼎盛时期。他们的世界观已确立，职业已定向，精力充沛，是承担社会义务、显示创造力的第一个阶段，并富有开拓精神，能很快地吸收新的思想或新的科技成果，并迅速地运用于工作之中。在参加考试的动机方面，一般有下列特点：

（1）显示力量和学术功底，希望社会给予正确评价；使所取得的成绩能得到社会承认，以便在工作中发挥作用。

（2）通过考试取得某级或某种合格证书，以便提高自身的社会或经济地位。

（3）少数人试图通过考试取得某种合格证书而变换工作职位；也有极少数人是出于碰运气的思想参加考试，能否取得某种荣誉或资格，无关紧要。

2．成年期。（女性40～55岁，男性40～60岁）

成人成年期的生理机能和心理功能，将从鼎盛的高峰逐渐向衰退方向迁移，但精力仍然充沛。在心理方面，机械识记下降快，意义识记虽有消退但不明显，聚合思维或发散思维有所发展，感情由热情转向深沉，意志坚强，兴趣由广泛迁移到专注。世界观更为坚定，对专业的态度更加炽热。对社会的责任感和义务感与关注子孙后代的幸福凝成一体。因此，是事业成功和创造力量显示的第二阶段。干事稳键，开拓性或改革性的工作只要看准

后，就能坚定不移地去实践。他们参加考试常常出于下列动机：

（1）专业上的进取，创造性成果鉴定的需要。

（2）职称评审或晋级的需要。

（3）荣誉或经济上的需要。

3．老年期。（女性55岁以上，男性60岁以上）

成人的老年期是指退休以后的岁月。其生理机能和心理功能衰退反映明显，精力没有前段充沛，在某些方面感到力不从心，但社会责任感和事业心反而更强，感觉时间紧迫，希望为子孙后代或后继者留下更多的业绩。随着时间的迁移，往往出现完美、遗憾、忏悔及"烈士暮年，壮心不已"的不同情感。老年参加考试的动机，以享受为主，满足心理需要，或弥补青壮年时代之不足。

（二）成人各年龄段考试的原则和方法

成人因年龄段不同，身心特点各异，考试的目的有别，考试的原则与方法也不尽一样。

成年早期与成年期成人的考试原则与方法，因成人的成年早期和成年期特点相近，考试目的相同，故考试原则和方法大体一致。成人考试的基本原则是：

1．定向性原则。

成人职业已确定，考试应以成人所从事的专业为主。但由于科学技术迅猛发展，新学科和新专业层出不穷，因此考试不能排斥新学科或新专业的内容，使考试能促进专业上的深化，积极推动新学科、新专业的发展，而不是为取得学历开绿灯。

2．创造性原则。

成人在这个阶段，是作贡献、出成果、创造发明的阶段，考试应激励他们多出成果、多做贡献，为社会建功立业。同时，考试应鼓励他们创新。对他们的创造性成果或开拓性项目，要在认真检验的基础上予以承认。

3．补偿性原则。

人的发展是不平衡的。青年时代应掌握的知识和技术，有些人因某种原因没有掌握，而在成年阶段通过自学或实践已经掌握，应允许他们参加补考。凡合乎规定标准者，应颁发合格证书，或承认某种资格。

4．求实性原则。

成人考试的立足点是提高广大公民的政治素质和业务能力，促进社会生产率的提高。因此，考试不能脱离学习与工作的实际，或脱离社会的客观需要。

现行教育和社会两大系统的成人考试，可概括为三类，即合格考试、选拔考试和成就考试。合格考试的方法，以口试、笔试及实际操作为主，或以专题设计与专题论文及答辩相结合的形式进行。选拔考试的测试方法以限时性口头回答或书面回答为主，也可采用变通式初试、复试及终试的方法进行考试。成就考试一般是根据成果进行答辩及评议，也可采用实际操作或演示的方法。总之，方法是根据实际情况和需要而确定，没有固定的模式，以有利于实现预定的考试目的为原则。

成人老年期的考试，应以鼓励安慰为主，方法应灵活、多样，具有娱乐性质。切忌超负荷量的考试，以防因过度兴奋或紧张而发生意外。

第二节　考生身心特殊差异与考试

考生身心特殊差异者系指非常态考生，其中包括超智考生、特殊才能考生以及身残考生。由于非常态考生的身心发展有别于常态考生，需有特殊的考试形式、原则和方法与之相适应，所以对非常态考生的身心特点与考试的相互关系辟专节进行探讨。现将非常态考生的考试原则和方法，按超智考生、特殊才能

考生和身残考生三种类型，分别阐述于后。

一、超智考生与考试

超智考生是指天赋优异、才智超常的儿童或少年考生。我国俗称为神童，西方则称为天才儿童。古时对超智儿童、少年的发现和教育已有论述。我国古代教育家孔子有上智之说，孟子有教育英才之乐等等；在西方，古希腊柏拉图首创“天才”一词，并对才智超常者有所论及，至于系统研究超常儿童问题，西方直到19世纪下半叶高尔顿（1822～1911年，英国实验心理学先驱，差异心理学的创始人）才首肇其端。他先以调查的方法研究天才与特殊能力者，继而创办实验室，利用仪器对人进行心理测验，研究人的个别差异。继高尔顿之后，研究者不断增多，成效显著的有司腾、比纳、推孟等人。随着社会的进步和科学技术的发展，社会对超智儿童和少年的发现、培养日益关注，为培养杰出人才相继出现了英才教育、超智儿童教育。心理学家和教育学家，也对超智儿童和少年进行广泛研究，既有超智儿童和少年身心发展特点问题的探讨，也有培养、鉴别和选拔方面的研究。下面就如何根据超智儿童和少年的身心特点设考与施考的问题作进一步的探讨。

（一）超智考生的身心特点

根据科学工作者的调查，超智儿童或少年在同龄人中约占千分之三。其生理、心理的发展较之同龄人有其明显的特点和优势。在生理方面，外国科学工作者曾对549个常态儿童和超智儿童运用37种精确的测量，作过22000人次的检查，结果表明，超智儿童不但在身高与体重上优于常态儿童，而且在各种生理品质上比常态儿童好，如握力、肺活量比常态儿童大，臂部、腰部、肩部各种阔度比常态儿童好。1980年，中国科技大学按照国家体委、原教育部、卫生部颁发的体质测量标准，对全体少年班大学

生进行了身体、形态、机能、素质等方面共23个项目的全面测试，结果证明，他们新陈代谢功能强、效率高，体格健全，以及动作协调、灵敏度、时间的掌握等功能，也相等或优胜于其他少年。他们在大学尽管学习任务繁重，但仍可抽出时间参加有组织的游戏。以充沛的精力从事体育活动，并能顺利通过国家颁布的青少年体育锻炼标准。由上可见，第一，超智儿童、少年身体发育正常，其生理素质与常态儿童少年相等或较优。第二，超智儿童、少年身体素质能胜任繁重的学习任务，他们也能够承担较常态儿童少年负荷更大的考试。

在心理发展方面，超智儿童、少年的共同特点是：

1．感知细致深刻、观察敏锐。

超智儿童、少年感知的整体性、选择性、恒常性以及有理解自已的规律，表现其视觉、听觉辨别能力发展突出，在感知过程中能抓住事物的整体，观察中能迅速的作出选择，可发现一般儿童觉察不到的情况或细节，能辨别常态儿童不能区别的东西，抓住事物的特点。例如，长沙市有位4岁的女孩参加小学一年级拼音统考得98分，她对汉字的偏旁部首、间架结构、听写极少错误，说明该生听觉、视觉能力高度发展，感知细致、深刻，观察敏锐，反映迅速。

2．注意集中，记忆特强。

超智儿童、少年在活动时，大脑皮层某一部位常形成一个强大的占优势的"兴奋灶"，而皮层的其他部位却处于抑制状态，表现出注意力高度集中，常以充沛的精力和充分时间解决一个问题，如听课、读书、解题、写作等等。这种状况一般可坚持2～3小时，他们的学习效果和工作效率等，都超过常态儿童和少年。

超智儿童、少年常有博闻强记、过目成诵的特点，记忆特强。他们识记不仅速度和持久性在同龄人之上，而且善于发现事物的规律，以自已独特的识记方法进行记忆。超智儿童、少年对

外界的刺激，由于感知的选择性，在大脑皮层形成有意识记的暂时联系，并加以强化使其持久，在需要时可使识记的事物再现，这就是超龄儿童、少年记忆的特殊性。

3．思维敏捷，想象丰富，勇于独创。

思想敏捷，善于捕捉问题的关键进行分析、概括、判断，是超智儿童、少年突出的特点之一。如解决数学问题能简化运算步骤。这是抽象逻辑思维高度发展的具体反映。

思想通畅，善于举一反三，提出非同凡响的主张，思考问题有与众不同的思路和方法，写作别有新意，是超智儿童、少年求异思维迅速发展的又一特点。这是由于超智儿童、少年在思维过程中，第一、二信号系统能协同活动以及第二信号系统——言语活动为主的分析综合能力超于常态儿童、少年的结果。

超智儿童、少年一般想象丰富，联想甚多，幻想奇特。他们大脑中新的观点或形象来自合理的联想、幻想或顿悟，而且具有独创性，表现出独出心裁的见解或新颖的形象，以及勇于弃旧图新，别开生面地解决问题或进行创造和革新的活动。

4．求知强烈，勤奋自信，富有进取心。

超智儿童、少年都有好学、好问、好奇的特点，对现成的事情不盲从，乐于追根溯源。他们勤奋、自信，学而不倦，锲而不舍，自已认为有价值的事情，即使遇到阻挠或利诱，也不轻易改变其信念，为达到目的而坚持拼博，直到问题解决或事情完成才罢休。

5．自我控制能力较强。

超智儿童、少年有强烈的自我意识，能较好地控制个人情感，很少计较个人得失。

超智儿童、少年在素质上虽有其先天优异之处，但关键是后天的早期教育才使其得到发展。

（二）超智考生的考试原则和方法

由于超智儿童、少年在身心发展上的特殊，他们在生活和学

习中必然有别致的表现。因此，对超智儿童、少年需有正确的认识，应给以恰当引导和培养，以便发挥他们的优势。在现实生活中，由于人们认识的局限性或偏见，常将学习成绩优劣做为衡量超智儿童、少年的唯一标准，而所谓成绩好，又仅指记忆书本知识的多寡，很少注意到其他智力和非智力因素，以致许多超智儿童、少年因未被发现，得不到正确的培养而被断送。发现、选拔超智儿童、少年，首先需要有正确的人才观和教育观，对超智儿童、少年的含意有所了解，然后按照超智儿童、少年的标准进行客观的考试和选拔，而不是带有偏见的主观臆断。

超智儿童、少年考试的原则和方法既是发现选拔超智儿童、少年的考试原则和方法，也是超智儿童、少年平时测试的原则和方法。

1．超智考生考试的原则。

（1）全面测试的原则。即坚持德、智、体、美、劳全面考核，不仅扬其长，而且要揭其短。

全面考核考生德、智、体、美、劳诸方面，是为了解他们各方面的发展水平。过去往往只重视智育的测试，而忽视德、体、美、劳的考核，因而产生一美盖百丑的现象。超智考生有才无德的现象并不少见，少数对社会缺乏责任感，通常表现出唯我至上的特殊优越感，缺乏独立生活、自理的习惯和能力，甚至追求享乐。这样的超智儿童、少年，非但成不了国家的栋梁之才，反而会成为社会的祸害或包袱。通过全面测试了解考生，有利于长善救失，使其健康成长。

（2）督导原则。人们对超智考生总是寄以希望，期待他们成为国家、民族的栋梁之才。因此，对他们的考试应是鼓励鞭策他们上进，引导他们向正确方向发展，决不能挫伤其积极性，压制其才智的发展，尤其不能伤害其自尊心和自信心。因为一般超智考生自尊心和自信心特别强，好胜敢拼，若挫伤其锐气，无异

于切断他们前进动力的能源。基于上因，考试后必须做好他们的善后工作，对考试胜利者要提出更高的要求，鞭策他们更上一层楼；对考试失败者则要进行诱导，帮助他们分析失败的原因，鼓励他们崛起，直至胜利。

（3）知识和能力并重的原则。超智考生在发展过程中，虽然天赋优异，但由于环境和教育之故，他们的发展，有些基础知识厚实牢固，是循序渐进上升的，有的则是跳跃式前进的，或超越了某些阶段，或发展有所偏颇，对某些学科掌握较好，而某些学科因接触少或无人指导，往往掌握一般；有些考生的知识面广，而且有一定的深度，考试可对答如流，但动手能力欠缺。因此，对超智考生的考试，只有坚持知识和能力并重的原则，才能促进超智考生发展成为出类拔萃的人才。

（4）因材施测的原则。"金无足赤，人无完人"，对超智考生反对求全的观点，同时也反对一美盖百丑的片面意识。超智考生的考试应按标准从严要求，但超智考生处于发展之中，他们不可能完美无缺，求全责备同样不可取。然而，也不能因超智考生有某种优势，对其缺陷采取容忍或掩盖的态度。这些观点或态度都不利于超智考生的发展。超智考生因年幼，以及所处环境与接受教育的不同，考试应因材施测，不宜采用整齐划一的统考形式。当然，已被选拔的考生在经过教育培养之后，可以应用统考的形式，但试题应有一定弹性，以发挥考生的优势，促使杰出者脱颖而出。

2．超智考生考试的方法。

超智考生考试的方式和方法必须谨慎选择，既要符合超智考生考试的原则，又要适应考试目的和内容的需要。超智考生考试如果从考试目的进行分类，基本上可归纳为选拔考试、合格考试和成就考试。三类考试因目的不同，方法也各异。

超智考生的选拔考试，是淘汰性质的考试。一种是发现超智儿童、少年的甄别性选拔考试，另一种是超智儿童、少年竞争性

的选拔考试。前者的目的是从常态儿童、少年中选拔超智儿童、少年，其依据是超智儿童、少年的标准。其方法多数是个别考试，以口头质问为主，当然也有当面笔试，或实际演示。后者是在超智儿童、少年中优中选优，其方法一般是团体考试，采用标准化考试或论文式考试，甚至变通式考试。口试、笔试可兼用。

超智考生的合格考试，是超智儿童、少年在某级学校、某种专业、某一学科学习中达到相应合格标准的认可考试或资格考试。由于超智儿童少年的特殊性，既可参加集体考试，也可进行个别考试，一般采用的方法是目标参考性考试和论文式考试。其测试方法既可用笔试，也可用口试，甚至采取设计或论文答辩的形式进行。

超智考生的成就考试，是检测超智儿童、少年成就和才华的考试。考生将创造发明的成果、科研的设计或论文以及创作呈献于评审委员会，接受鉴定及评议。一般形式是先由考生宣读设计或论文，再进行答辩，然后由评审委员会作出鉴定。

二、特殊才能考生与考试

特殊才能的考生，是指有特殊才能的儿童、少年。他们的素质和智能与常态儿童、少年相似，但在某一方面有独特才华，如绘画、书法、音乐舞蹈、戏剧表演、体育、棋艺、文艺创作、科技创造等不同方面有优异表现或创造力。其特殊才能一般是由家庭的熏陶、自身的特殊爱好、或学校培养而成。

特殊才能的儿童和少年，被各国视为国家的宝贵财富，对他们的选拔和培养至为重视，常以各种方式对他们进行培养。

（一）特殊才能考生的特点

1．有特殊的爱好和兴趣。

特殊才能的考生由于家庭的感染或教师的培养，逐步形成对某项活动的特殊感情与爱好。随着爱好或兴趣的发展，逐步成为

一种专一炽热的职业心理，使其达到忘我迷恋的程度。同时，他们感觉器官的视觉、听觉、触觉等能力也向其爱好的专业方向发展。如有绘画特殊才能的儿童、少年，对颜色的辨别能力远远超过常态儿童、少年。

2．特殊才能与认识过程两者发展密切关联、相互促进。

具有艺术特殊才能的儿童、少年，观察力注意力偏重于声、色、形象的刺激，能敏锐地捕捉住爱好所需要的焦点。因此，记忆能在大脑中再现观察的形象。其想象、思维都是形象性的，创造性活动也具有形象特征。又如具有数学特殊才能的儿童、少年演算或证明数学题的注意力，一般集中2～3小时，抽象的逻辑思维能力特别发达，分析、综合、判断概括有条不紊，其思维的开阔性、通畅性及完整性为一般人所不及。

3．情感深厚、稳定，意志坚韧、顽强。

特殊才能的儿童、少年情感深厚、稳定、表现在对某项专业的专注上，不见异思迁或时冷时热。当某项活动获得成功时，常常会暴发出奔放的激情。他们的社会情感——道德感、美感、理智感的发展及成熟期，要早于常态儿童、少年，这是他们特殊才能发展的先导，原因在于他们对爱好的专业或专长的体验深度超过一般人。同时也是特殊才能的儿童、少年意志坚韧、顽强动因所在，在遇到困难或外因诱惑时，他们仍能拼博或排除干扰，竭力追求成功的目标。

（二）特殊才能考生的考试原则与方法

第一，特殊才能考生一般在某一方面具有超常的才华，而其他方面却显得平凡，甚至很差。因此，特殊才能考生考试的首要原则是全面测试，全面评价。既要了解超常才华的深度与广度，也要了解德、智、体、美发展的水平。只有全面测试、全面评价，才能高屋建瓴，全面施教，培养出卓著的专家，而不是匠人。

第二，先天遗传素质只是人的发展的物质前提，后天的环境

和教育才是决定性的条件。由于人的先天遗传素质与后天环境教育有差别，因而人的个别差异客观存在。特殊才能是人的个别差异之一，与其生理素质、环境、教育密切相关。对特殊才能考生的考试，不仅要检测考生的特殊才能表现，而且对考生与专业相关的生理和心理素质也要进行检查。

第三，标准统一。特殊才能考生必有超常才华的表现，如书法、美术的作品——画、雕塑，以及诗歌、舞蹈、戏剧的表演等。由于人们对作品及其创造和表演的体验不尽相同，评价也难免有分歧，在测试前必须统一测试和评价的标准，以免出现因评价有争议而贻误特殊人才培养的现象。

第四，特殊才能考生才华各殊，考试要因人而异，个别测试是考试的主要方式。考试方法既可笔试，亦可口试，但必须以演示测试为主。

三、身残考生与考试

身残考生是指盲、聋、哑及肢体残疾者。这部分人中蕴藏着巨大的能量，他们身残志不残，通过艰苦拼搏，往往能成为自食其力的劳动者甚至能成为作家、诗人、科学家、医生等等，可为社会作出巨大的贡献，理应受到社会的尊重与爱护。当然，在身残者中对生活失去信心、缺乏生活能力的人是存在的，关键的问题是社会应本着人道主义精神，对他们关心、尊重和爱护，使他们热爱生活、追求生活、不断增强生活的信心与能力。现在，世界各国都十分重视身残者的培养，为他们积极创造条件，让其尽才能于社会。在我国，党和政府对身残者给予巨大的关怀。据1986年统计，专为盲、聋哑者开办学校387所，在校学生40000余人。又如将张海迪、史光柱等身残人才送入高等学校深造，使他们能更有效地发挥其才能。

由于身残者是社会的一支重要力量，考试作为社会开发、利

用人力资源的手段，为身残考生积极拓宽道路，是社会赋予的责任。因此，对身残者的考试问题必须研究和探索。

（一）身残考生的特点

其一，身残考生虽然生理上存在缺陷，但人是有机体，身体的某一器官受到损伤，其他器官通常能给以补偿。如盲人虽然视觉受到损害，但听觉和触觉特别发达，可补偿视觉的功能，借助触觉和听觉阅读盲文；聋哑者听觉受到损伤，失去语言能力，但视觉和动觉发达，可借以补偿听觉和语言能力，如聋哑人的手语就是他们交流思想和学习的工具。

其二，身残考生在生理上有所缺陷，但对生活具有炽热的情感，自尊、自信、顽强是他们生命的支柱，渴望理解，需要尊重，要求自立。他们迫切要求学习，希望具有独立生活能力。所以，他们在学习和工作方面的进取心往往超过常人。

其三，身残考生的心理过程由于生活环境和生理条件的原因，一般形成了较良好的心理品质，记忆持久而准确，注意力集中而不易被干扰，思维灵活而敏捷；情感稳定而有自我控制能力，意志坚韧而具有自觉性。

身残考生身心的特点独具，组织考试应按照考生的特点进行安排。

（二）身残考生的考试原则和方法

首先，身残考生在生理上有所缺陷，他们在心理上非常敏感，需要人们理解。对他们来说，最重要的是尊重和爱护，而不是可怜和同情，考试要激励他们奋发进取，自信、自尊、自立。既尊重考生人格，又能严格掌握标准，这是考试的基本原则。

其次，身残考生情况特殊，应个别考试与团体考试相结合，从鼓励引导出发，从易到难，从简单到复杂。盲文、哑语是盲、聋哑人语言工具，应正确使用，教授与考试要紧密结合，不能两者脱节，或因考试失误，使他们心灵上受到损害。

再次，身残考生的学习都有其职业上的专业性，考试必须围绕专业进行。如果超越其范围，不仅达不到测试的目的，反而会产生不良影响。其原因是：身残考生一是敏感，二是自尊心强，甚至近乎固执，如果组织考试不慎重，就会产生负作用。

最后，身残考生生理上虽有缺陷，但又有某些特殊功能，如盲人耳必聪，触觉灵敏；聋哑者眼必明，手非常灵巧。其他身残考生也是如此，考试设计应将他们所学专业与生理特长结合起来，使其优势得到充分发挥。

第三节　考生家庭、社会环境与考试

“人们的社会存在决定人们的意识”，①这是历史唯物主义的一个根本观点。考生对考试的态度和行动无疑为社会所决定，具体地说是受家庭和社会环境所影响，但世上的事物都是由各方面的因素决定的，关系极为复杂。深入分析考生的家庭、社会环境与考试的关联，从不同的方位揭示考试的规律，将有益于考试的实践。

一、考生家庭环境与考试

家庭是社会的基本单位，它的经济状况、政治地位、文化水平、职业及人口结构等，都是影响考生考试态度和行动的因素。

经济是社会的基础，决定着社会的政治状况和文化教育及其他事业发展的规模和速度。同理，家庭的基础不仅决定家庭其他方面，而且支配着家长和考生对考试的态度和行动。如美国的小康之家的子弟都争取到社区学院去学习，虽然有能力考取名牌大

①《马克思恩格斯选集》第2卷，人民出版社1972年第1版，第82页。

学，但经济上无力支付学习和生活费用，只好忍痛割爱。又如印度、阿拉伯国家的豪门巨富，他们不惜重金送子女到欧美去投考中学、大学，甚至小学，而不管其智能状况如何或成才与否。在日本，有产家庭则要求子女投考名牌小学、中学、大学，参加应试战争，以期望子女挤入高层社会。

家庭的政治地位和态度，同样是影响考生考试的因素。如英国贵族要求子女从小报考名牌私立学校，而大学则以考入牛津、剑桥大学为奋斗目标，因为牛津、剑桥不仅是科学家的摇篮，更重要的是政治家的培养所，英国历史上有29位首相都是出身这两所大学。若挤进牛津、剑桥，便身价百倍，无疑在政界增强了竞争力。在我国革命史上，有不少革命先辈出身于剥削阶级家庭，他们的家长则希望他们成为本阶级的继承人。他们出去投考求学则背叛家庭投身革命，成为剥削阶级的掘墓人，如澎湃烈士等等。这说明家庭的政治地位，是影响考生考试的重要因素，但决定因素是考生本人，能否成才？走什么道路？都取决于考生自己。

家庭的职业与文化素养也同样影响考生。音乐世家的子女，从小就受音乐的熏陶，便喜爱音乐，促使他们投考音乐学校。英国高尔顿调查了30家有艺术能力的家庭，发现他们子女的64%有艺术能力；而150家无艺术能力的家庭，他们的子女就只有21%有艺术能力。家庭文化素养不同，家长和考生对考试的态度也有异。现阶段，我国农村实行家庭承包责任制后，农村经济迅速发展，农民温饱问题基本解决，有能力负担子女读书。但对考试却出现两种显然不同的态度：文化水平高的家长，大力进行智力投资，要求子女努力学习，积极参加考试，争取进入高一级学校深造，文化水平低的家长，除少数因无文化之苦而督促子女好学上进外，对子女学习一般采取漠不关心的态度，对子女考试成绩的优劣也无所要求，甚至要求子女辍学做帮手。当然前一类家长要求子女严，子女必然以积极态度对待考试，而后一类家长对子女无

所要求，必然使子女对考试采取消极应付的态度。

家庭的人口结构，也是影响考生考试的因素之一。家庭人口少，经济状况良好，家长有条件指导子女学习，也必然以积极的态度关心子女考试，子女也会认真对待考试。如果家庭人口虽少，但经济状况不佳，整日为生活而奔波，也就无力关心子女的学习和考试。这种家庭的子女，他们对考试的态度多决定于学校的教育。家庭人口多，对考生考试产生两种不同的影响（排出经济因素影响），若长子长女学习积极，考试成绩优异，则对幼子幼女有好的影响；若长子长女学习成绩差，对考试消极应付，也有可能影响幼子幼女考试的态度，一般是消极的多。当然也不排除有少数积极者。如我国一个失去配偶的普通农村妇女，要将三个子女培养为高级专门人才，在经济条件并不宽裕，她也不能对子女进行辅导的情况下，虽全靠社会创造条件，但家庭教育，兄弟之间的互相影响，仍然是他们成才的重要因素。

家庭环境对考生的影响是多因素的，核心是经济基础。在经济状况相同的条件下，政治地位又变成了主要因素，不过文化素养与职业也在某种情况下变为主要因素起作用。家庭人口结构，对考生影响既可起积极作用，也可起消极作用。总之，考生家庭环境对考生考试的影响是客观的，切不可忽视家庭环境与考试的关系。

综上所述，家庭、考生对待考试有三种不同的态度：

第一种，努力备考，奋力拼搏，积极竞争，用考试来检测知识和能力的水平。不以分数高而喜，也不以分数低而忧，而是立足于成才，在考试之后，进行分析综合，积极准备再次考试，以攀登新的高峰。

第二种，考试至上。如日本多数家长认为参加考试尤其是升学考试是决定命运之战。考生除在普通学校正常学习外，还在课外参加补习学校，致使考生疲于奔命，整天处于紧张状态。到考试时，临场不能正常发挥。考试成功者欢欣若狂，失败者痛不欲

生。这种对考试的态度正是学力社会中考试的一种病态。

第三种，无关紧要。家长对考试抱无所谓的态度，考生则应付了差。反映家长对子女无所要求，考生也无成才的理想。关键问题是家长不明确考试的意义，看轻自己对子女成才所负的责任，不能把考试视为激励和鞭策子女上进的一种措施。

二、考生社会环境与考试

考生社会环境与考试，系指考生所处社会的政治、经济、文化科学、教育、考试、传统意识诸因素对他的影响，以及在这种环境影响下考生对考试的心理状态和考试的关系，亦即社会环境考生、考试三者之间的复合关系。

人的思想意识是现实世界客观事物及其过程的抽象反映。从时间而言，一般是局限于当代社会这个历史范畴，超前或滞后的观念只是特殊现象，但并不排斥社会思想意识的继承性。考生对考试的心理状态，也是当代社会意识的反映。从空间而言，人总是生活在一定的社会关系之中，它的本质是一切社会关系的总和。在阶级社会中，人的思想意识无不打上阶级的烙印，是受具体的阶级关系所制约和决定的。在没有阶级的社会中，人的思想也不能超越他所生活的社会环境。一个人的思想意识只能是他在社会生产关系中所处地位的反映。当然，考生对考试的态度，也要受到他在生产关系中所处地位的制约。但是，世界上事物是复杂的，矛盾是多因素的，决定事物发展是主要的矛盾或主要的矛盾方面。考生的心理也是如此，外因必须通过内因起作用。在现代社会中，生活有保障的考生存在心理状态一般表现为：其一，积极进取，不安于现状，追求新科学、新技术、新生产部门的开拓。因此，他们通常报考社会发展急需的专业，如有关电子、太空等尖端专业，并积极备考，不懈地为之拼博、奋斗。其二，满足于现实的物质生活，追求安逸舒适的生活环境，他们向往

的是没有风险又不付出艰辛劳动的行业，如银行、商业部门等。因此，他们大都投考财经专业，如美国的青年投考财贸的多，报考理工的少，现今美国科技界，美籍外裔特别是亚裔科研人员已占1/3。其三，无所追求，安于享乐，靠父母权势或财富生活，对考试是应付差事，甚至以鄙薄的态度注视别人去参加考试。在西方世界，不管是发达国家，还是发展中国家，社会财富占有和分配是不合理的，贫困失业成为社会突出的问题、广大的群众贫困、失业，迫使他们追求果腹。因此，不管是普通劳动者，还是有专业的知识分子，都去参加竞争的招工考试，其心理上唯一的希望是争取被录用。而无生活保障的青年学生，报考职业学校或高等学校，首先考虑的是将来就业，而不是自己的兴趣和爱好。其次，目标确定后是去苦斗，参加竞争。他们的考试态度既严肃认真而又忐忑不安。

人都在一定的社会政治环境中生活，无疑社会政治对他产生不同的影响。对政治的态度，有积极拥护者，有消极应世者，也有批判反对者。由于考试是为政治服务的，人们对它的态度有积极的参加者，也有消极的应付者，还有抵制者。中国封建社会延续到清代已腐朽没落，清王朝用科举考试来维护摇摇欲坠的封建制度，人们对科举考试的态度是泾渭分明。吴敬梓的《儒林外史》就是历史的见证。作者对科举考试进行了无情的揭露和深刻的批判。书中对王冕、杜少卿、荆元轻功名富贵，品行端方，清廉高洁，笑傲王侯，才高学深进行了赞颂。对周进、王惠、严贡生、范进、匡超等人热中举业，醉心功名富贵，积极拥护科举进行了无情地鞭挞。对维持文运的虞育德、庄绍光、迟衡山等人看轻举业，要求自食其力，轻视科举的态度给予肯定。

从15世纪末至19世纪中叶的300年中，英国不仅完成了资产阶级革命，而且兴起了并实现了工业革命，成为世界上头号资本主义国家。资产阶级运用考试手段为其选拔和培养人才。

如争夺海上霸权，进行殖民掠夺，发展舰队和商船，英国政府大力培养与选拔海事人员，并规定王子与青年贵族必须到海军服务。在此种政治环境下，英国青年纷纷投考海军与海军学校和海事学校，以参加海军为光荣，以被录取到东印公司为骄傲。

文化教育的发达与否，标志一个国家或民族的进步与发展程度，当然各国统治者也是重视文化教育的发展，不过其目的是为本阶级服务。如欧洲的中世纪是宗教统治文化教育的时代，甚至宣称科学是上帝的奴婢，学校掌握在教会手中，虽然世俗的人文主义先驱者们反对宗教文化教育，但事实上是宗教文化教育占统治地位，考试也是宗教文化教育的附属物，人们以推荐与录取神学院为荣，以当上神职人员为幸事。

在当今世界文化教育向世界化方向发展、外语成为获取和更新知识的重要手段的情况下，青年除学习母语外，积极学习外国语，投考外语专业成为热门。汉语是世界各种语言中比较难学的，但今天世界大多数国家除有汉学研究单位外，大学还开设汉语专业，连中学也开设汉语课程，还有些人利用假期到中国来学习汉语。可见，文化，教育发展的倾向性能给人们的报考方向以重要影响。又如当管理科学兴起时，青年及企业人员便掀起学习管理科学的热潮，高校管理专业成为投考最多的热门专业。

文化教育发展的不平衡，也会导致学习与投考倾向的不同。就我国而言，上海、北京的高中毕业生愿投报本市中专或本市的高校。湖南、湖北的高中毕业生则向往全国著名大学，如北大、清华、上海交大、复旦等。广东青年倾向投考外语、化学、生物、体育等专业。

再者，任何国家或民族的传统习惯，都影响着人们的行为倾向。如东亚某些佛教盛行的国家，青年必须到寺院出家一段时间，受佛教教育。日本经济发达，学力风气盛行，青年一代拼命争取考入名牌学校，社会侧重学力而不问其能力。我国长期受儒

家思想的影响，在人们的思想意识中，修身齐家治国平天下的影响仍然存在，追求理想、好学、重义轻利，时刻不忘党和人民。如许多在国外学成者奉"我的事业在祖国"之念，坚定地回国，立志为祖国服务。这是民族意识的结晶，也是民族传统习惯给人民的影响所致。

文化教育的发达与否，标志一个国家或民族的[illegible]度，当然各国统治者也是重视文化教育的发展，不过其目的是为本阶级服务。如欧洲的中世纪是宗教统治文化教育的时代，甚至宣称科学是上帝的奴婢，学校掌握在教会手中。虽然世俗的人文主义先驱者们反对宗教文化教育，但事实上是宗教文化教育占统治地位。考试也是宗教文化教育的附属物。人们以推荐与录取神学院为荣，以当上神职人员为幸事。

在当今世界文化教育向世界化方向发展，外语成为获取和更新知识的重要手段的情况下，青年除学习母语外，积极学习外国语，投考外语专业成为热门。汉语是世界各种语言中比较难学的，但今天世界大多数国家除有汉学研究单位外，大学还开设汉语专业，连中学也开设汉语课程，还有些人利用假期到中国来学习汉语。可见，文化、教育发展的倾向性能给人们的报考方向以重要影响。又如当管理科学兴起时，青年及企业人员便掀起学习管理科学的热潮，高校管理专业成为投考最多的热门专业。

文化教育发展的不平衡，也会导致学习与投考倾向的不同。就我国而言，上海、北京的高中毕业生愿报考本市中专或本市的高校。湖南、湖北的高中毕业生则向往全国著名大学，如北大、清华、上海交大、复旦等。广东青年倾向报考外语、化学、生物、体育等专业。

再者，任何国家或民族的传统习惯，都影响着人们的行为倾向。如东亚某些佛教盛行的国家，青年必须到寺院出家一段时间，受佛教教育。日本经济发达，学力风气盛行，青年一代拼命争取考入名牌学校，社会偏重学力而不问其能力。我国长期受儒

第六章

考试的类别、方法与原则

社会分工及其对人的素质与智能要求的特殊性，决定了考试门类的多样性。而考试对象和测试要素的共同性，又将因分工划分成的众多考试门类，归并为相对稳定的几种类别。

人的身心个别差异的不同方面及其外化形式的特殊性，决定了测度、甄别人的德、学、才、识、体个别差异之考试方法的多样性。而科学性、客观性、公平性、实践性又是人类各种考试实现预期目的必须遵循的基本原则。

考试这一社会现象有其自身演变、发展的规律，它始终受社会发展和人的发展规律所制约。考试的具体目的、内容、性质、门类、方式、方法与手段，总是与社会生产力和文化科学技术的发展水平、社会生产和社会生活对人的素质与智能的要求，以及人们对考试的使用范围和认识程度相一致。一般说来，考试的具体门类，是随社会分工及其对劳动者素质与智能要求的变化而变化的，而考试的方法则取决于人的素质与智能外化的形式。多样性是考试门类与方法的基本特点。

但是，为考试的本质属性和职能所决定，人类社会不同历史发展阶段及社会各个部门的考试，又有许多共同的特征：

第一，一切考试的对象都是人；

第二，人的德、学、才、识、体诸方面的特质，是构成各种考试客体的基本要素；

第三，测度、甄别应试者德、学、才、识、体某一方面或各个方面质与量的差异，是人类所有考试的本质目的，只是对德、学、才、识、体测试方面的多寡，以及同一方面测试的具体内容与标准等有所区别而已；

第四，不同个体同一方面的素质与智能，其外化形式不无同一性，可用相同的方法进行测度或甄别。

在长期的考试实践中，人们为使考试的组织和实施简便易行，便根据考试自身的规律，将因分工划分成的众多考试门类，以及为人的素质与智能外化形式的特殊性所规定的多种考试方法，从考试的内在结构与外部特征的共性方面，逐步归并成相对稳定的几种考试类别和方法。

第一节 考试类别

尽管人们对人类社会各种考试作了大致的归类，但由于划分考试的角度不尽相同，现行考试仍然类别极繁、名称多样。为便于社会各行各业根据实际需要选择恰当的考试形式，我们从考试的方式、目的、要求、试题设计技巧、内容表现形式、施测形式、作答形式、分数反映形式、评估手段、隶属关系等十个方面，将现行考试作了相对归类，并对各类考试的适用范围、特征及功用分别进行论述。

一、按考试方式分

从考试方式的角度划分，现行考试可归为三类：面试、笔试与综合考试（面试、笔试并用）。

（一）面试

面试是人类社会最古老、使用最普遍、简便易行而又要求严格的一种考试。早在原始社会末期，面试就被广泛应用于社会生产与社会生活，后为历代沿用，迄今为止，它仍是现代考试的重要组成部分。

面试按其测试方式、适用范围与功用，又分为口试、操作考试、演示考试三种。

口试。它是一种主试者与被试者面对面的问答考试。在笔试产生之前，口试是学校检测学生学业成绩和社会其他行业考察人才学识与才能的主要形式。通常采用两种形式进行：一是个别口试。每个应试者按照考试规程依次接受主试者的检测，或通过随机抽签，就签上主试者事先拟定的题目作短时准备后，口头回答主试者的问题，或由主试者口头提出事先拟就的题目，应试者经短时思考后口头作答。至于现代考试中的论文答辩，则属个别口试的高级形式。它较之一般性口试要求更严，主要用于博士、硕士学业成就的综合性能力检验。论文答辩对应试者既有检验的一面，又有教育的一面，考试气氛严肃而庄重，主试者提问的范围以论文内容为限，而一般性口试的提问范围，则以主试者规定测试的内容为限。二是小组口试。应试者按规定数量组成应试小组，共同接受主试者的检测，考试以讨论式或问答式进行。小组成员既可分别回答主试者提出的不同问题，也可相互补充，即全组成员分别回答主试者提出的同一问题。此种形式为西方国家文官考试常用，也是某些社会团体举行知识竞赛活动时，采用较多的一种形式。

口试由于不受文字限制，主试者提问的范围可大可小、内容可深可浅（命题性口试例外），应试者在主试者所提问题的范围内，能广泛运用所学知识自由抒发见解，充分展现才能。所以，口试对应试者学识的广度与深度、口头表达能力、思辨能力、灵敏度、想象力、创造思维，以及性格和意识倾向等，能比较深入地检测或甄别。这是笔试所不及的。然而，由于口试是一种主试者与被试者面对面的对话或考试，应试者的心理往往处于紧张状态，加之备答时间仓促，应试者对所提问题缺乏深思熟虑的条件，对部分应试者来说，有受约束的一面，测试结果要准确反映应试者的真实水平较为困难。同时，口试多为个别测试，考试费时耗力，工作效率低。又因测试时间长，主试者易于疲乏，同一考试过程中的评分标准不易控制，前后效应性差，考试分数的误差成份大，直接影响考试结果的可靠性和有效性。因此，当笔试问世后，口试就在人类考试的发展史上由主要形式变为辅助形式。

操作考试。它是一种直接考试。这种考试是让应试者通过现场操作或具体的工作实践，直接检验应试者是否具备从事某项工作的知识、技能、技巧、素质与能力。操作考试历史远久，使用范围极广，如奴隶社会时期，奴隶主为检验青年的勇敢精神与征战本领而进行的夜间偷袭，以及捕杀奴隶的活动，封建社会民间“九老十八匠”各种技艺的出师考试，军事系统的实战演习、敌情侦察与兵器操作，交通部门各类运输工具驾驶执照的考试，公安部门的案情侦破与缉捕，司法机关的模拟法庭、教育系统的实验考试、教学实习，人事部门的干部试用，以及西方国家选拔文官时所进行的模拟性治事考试等，均属操作考试的范畴。

演示考试。它也是一种直接考试，即通过现场表演的方式，检验应试者某方面技艺掌握的熟练与准确程度，以判定应试者是否达到某种规定的标准。演示考试一般用于体育、音乐、美术、

工艺、等方面技能、技巧、技艺的检测与评定，以及机器操作与器械使用的测试。

（二）笔试

笔试是现代考试的主要形式。应试者书面回答主试者的问题，是笔试的基本特征。由于使用笔试的物质条件要求高于面试，需有较发达的文字，以及笔、纸、墨等，作为组织和实施考试的基础，所以人类使用笔试的历史要比面试短得多。中国在公元7世纪，笔试才取代了面试在考试中的主体地位，西方直到18世纪才由面试为主过渡到以笔试为主。

笔试的产生及其应用，是人类考试发展史上的一次飞跃，它适用于社会各系统的多种考试，具有四大优点：一是施测经济。无论是在单位、部门、系统范围组织的中、小考试，还是区域性、全国性的大规模考试，笔试都能在统一时限内以团体形式对所有应试者同时施测，可节约时间、人力与物力，降低考试的成本。二是内容深广。一张综合使用多种题型而又拼配合理的试卷，其试题大都在百题以上，容量较大，既能测试应试者认知水平的不同层次，也能顾及主试者欲测知识的各个方面，有利于多面、多层检测应试者的知识与能力。三是质量易控。笔试一般是团体性施测，所有应试者使用的是同一份试卷（或等值试卷），应试程序和规则相同，应试环境大体一致，有统一的评分标准，在施测过程中，对与考试无关的各种主客观因素的干扰能够有效控制，将考试的质量控制在规定的标准范围。四是考试结果比较客观。由于测试量尺、评分标准、实施程序与环境等，都有统一的标准和要求，且有考试过程中各个环节的误差控制，所以考试成绩一般能较为客观地反映应试者的实际。

但是，笔试的缺点也十分突出。它对应试者素质与智能的检测，除记忆性知识、记忆能力、文字表达、逻辑推理能力、独立判断能力等，是直接检测外，其它能力的检测多属间接，缺乏实

证性。若是传统考试的笔试，测试内容狭窄，评分易受评卷者水平、好恶、情绪的影响，以及卷面清洁、书写工整、字迹清秀等与测试内容无关因素的干扰，进而造成考试分数失真。此外，不易采用电子技术阅卷、评分、也是传统考试的测试不易克服的弱点。

笔试有开卷与闭卷考试之分，两者在检测应试者知识与能力方面各有所长。

闭卷考试的内容限于书本，多用于基本知识、基本概念、基本原理与技术知识等记忆性项目的检查。它有利于应试者全面系统地复习所学内容，深化和巩固已有知识。

开卷考试多用于创造性、综合性项目的测试，重在检查应试者掌握知识的质量及驾驭知识的能力，其内容涉及书本内外，应试者需综合运用所学知识才能正确回答主试者的问题。开卷考试只要命题灵活、严谨，应试者准备细致，作答认真，对于培养学生理论联系实际的学风，根治"填鸭式"教学和专注知识记背的弊端，提高后继人才的实际工作能力等，都有积极的促进作用。

（三）综合考试

综合考试是现代考试常用的一种测试形式。其基本特点是笔试、面试并用。采用综合考试的目的，在于全面检测应试者已有知识的程度、理论素质、学术水平、技艺水平和分析、解决实际问题的能力。它的使用范围不如笔试、面试广泛，一般在工矿企业的招工考试，部分高校系科专业的招生考试，某些特殊行业职工的晋级考试，以及大学一些实践性较强的学科考试时使用。综合考试可弥补单一笔试或面试的缺陷，能对应试者的素质与智能作较全面的检测，有利于准确甄别、评价、选拔人才。

二、按考试目的分

考试目的有广泛、狭义之分。狭义的考试目的，系指每次考

（或每一种考试）的具体目的，如中学毕业考试的目的，是检验中学生是否达到规定的培养标准。广义的考试目的，则指的是社会各部门所举行（具体目的有别）的不同考试的共性目的，如工厂招工考试的具体目的是选拔新工人，高等学校招生考试的具体目的是选拔新生，而选拔人才却是他们的共性目的。若从这种共性目的的角度划分，考试有六种不同的类型：

（一）　安置考试

这种考试主要用于学校教育对象的分流、分班或分组，社会各劳动部门新从业人员的分工，以及军队新战士的兵种分配。各级学校的新生、新入伍的战士及新录用的工人，尽管他们已经过考试或考核的筛选，但每个人的政治、业务与身体素质并不是整齐划一的，其素质与智能方面的个别差异客观存在。为使他们能在适合自己身心条件的环境中学习或工作，以便因材施教、量才用人，各单位往往要再举行一次安置性考试，如各级学校为恰当安置新生的班、组，对刚进校学生所进行的学前考试（又称摸底考试），工矿企业为合理分配新工人的职位而举行的就业前考试（又称工前考试）等，都属此类。这种考试无需统一的合格标准或选拔标准，重在鉴别应试者的实际水平与能力性向，以求能与职相称，教与学密切配合，从而促进人才培养效率和人力资源利用率的提高。

（二）成就考试

成就考试是一种检验应试者阶段性学习或训练效果的考试。测试内容限于应试者已学材料或训练项目的范围。这种考试既可由相同学科集体命题，也可由任课教师自行命题，但命题的范围和要求，都必须以教学大纲与教材内容为依据。按照测试的具体要求和内容，成就考试又分为两种：单科成就考试，即依学科分类分别测试学生各科成绩，如中学的语文、数学、外语等单科成就考试；综合成就考试，即在同一考试中包含多门学科的内

容，以测试文化素质或某种专业性的基础知识，多用于文化知识普查。这种考试应取材广泛，试题的难易要适度。各级学校内部举行的单元测验、期中考试、学期考试、学年考试，军事训练、体育训练过程中的阶段性测验，均属成就考试的范畴。成就考试的功用是，了解教学或训练的进度，检验一定阶段的教学或训练成效，为决定学生升留级和改进教学、训练内容与方法提供依据。

（三）合格考试

这种考试，又称达标考试、水平考试或证书考试。它是根据某种既定的统一标准，测量被试者的学业水平、技术水平或业务能力是否达到要求，以获取颁发某种学历或学位证书，以及授予某种专业技术职称的依据，如大、中小学的毕业考试，文、教、科、卫及其他业务技术系统的职称考试，运动员、裁判员不同技术等级的过关考试，各级体育锻炼标准的达标考试，部分职业的执照考试等。合格考试一般由主管部门或职能部门主持实施，以统一规定的标准作为判别应试者是否合格的尺度，而不以应试者测试的实际水平调节合格标准。这种考试的宗旨不是优中选优，它只论合格或达标与否。在现代考试中，合格考试是使用广泛、要求严格的考试之一。

（四）选拔考试

选拔考试是根据一定的目的与条件，从同类应试者中挑选具有接受某级教育或从事某种职业的所需人才。它的宗旨恰与合格考试相反，不是判定应试者的实际水平是否达标或合格，而是在同类应试者中选拔优秀者。因此，人们又将选拔考试称之为竞争考试或淘汰考试。它是一种要求极严、影响最大、适用范围极广的考试。选拔考试的标准根据实际需要确定，测试内容不限于应试者已学材料或受训内容的范围，主要依据应试者投考对象（某级学校教育或专业教育，某种职业）对欲选人才德、学、才、识、体诸方面的要求。反映成绩的方法一般是四种：百分制、等级制

（A、B、C、D、E……，或超、优、上、中、下、劣）、正态分制（标准分数、T 分数等）和积分制。按应试者成绩的高低序列依次择优选才，以录满预定数额为限，是选拔考试的基本原则。属于这类考试的有各类各级学校的招生考试，社会劳动部门的招工考试，事业、行政部门的招干考试，各军种、兵种的招兵考试，以及各类艺术人才的选优考试等。

（五）诊断考试

诊断考试通常在教学或训练的过程中进行，目的在于诊断一定时间内教学、训练的情况，通过测试所提供的反馈信息，分析教学、训练中存在的问题，了解教学（或训练）对象学习知识、掌握技艺的困难之所在，以及时改进教学、训练的内容与方法，不断提高教学、训练成效。诊断考试的测试内容，可根据教学大纲相应部分的要求和教材内容，由任课教师或主任教练自行确定。此种考试的测试结果，仅供教师教学参考，不作为评鉴学生学业成就的依据。所以，在性质与目的方面，诊断考试有别于平时测验、期中、期末等其它教学过程中的考试。

（六）研究性考试

所谓研究性考试，即是一种为获取从事某项研究的素材而举行的考试，属调查性质。其使用范围，一般限于教育、人事方面的研究部门或行政管理部门。如为研究教育质量升降的原因，探讨科学管理途径，分析比较不同时期教学大纲、教学计划、教学方法、教材内容的优劣所进行的考试，以及国际教育成就评价协会举办的各种标准化考试，便属此类。

三、按测试要求分

考试的测试要求因考试目的的不同而有异，大致可归为速度与深度两种要求。根据此种意义上的测试要求划分，考试可分为：

（一）速度考试

使用速度考试的目的，在于检验应试者掌握知识、技能的熟练程度与思维的灵敏度。例如，应试者在规定时间内究竟能回答（或完成）多少具有某级难度的测试题目（或项目），正误比率是否在规定范围。又如，应试者能否在规定时间内按要求完成所有的测试内容。因此，速度考试的时限从严，其试题或项目的份量，一般超过测试对象在规定时限内可能完成的负荷量。

（二）深度考试

深度考试又叫难度考试，多用于各种能力的测试。深度考试试卷题目的编排，通常是按照试题的难易度从低到高依次排列，意在检验不同水平的应试者能分别解答何等难度的题目，以区分同类应试者各自具有的实际能力。该类考试时限从宽，应让考生能在规定时间内答完全部测试项目（并不要求所有考生都能答对全部项目）。

但是，有些考试对速度与深度的要求并不是绝然分开的，常常是两种要求兼顾，既有能力的要求，又有时间的限制，象各种合格考试和选拔考试便是如此。它对速度和深度都有严格的规定。在命题、试卷编制和施测的过程中，均需体现上述两种要求，并从严控制。

四、按试题设计技巧分

考试试题的设计，既有严格的规定性，又有一定的灵活性。一方面，试题设计应符合两大规律：试题功能规律；人的素质与智能的外化规律。这是因为：一种试题只能对测试某一方面或几方面的知识、能力有效，而不是对各方面知识或能力的测试都有效；人的不同方面的素质与智能，各有其特殊的外化形式，它需有与之相适应的题式和内容，才能测到它所含的质与量。另一方面，对于测试同样内容之试题的设计技巧，又有相对的灵活性，

既可用一道含量较大的试题，去测试应试者某几方面的知识或能力，也可把某一方面的知识或能力进行分解，用几道或多道单项试题去测试应试者。这种试题设计的综合或分解的方法，便是试题设计的技巧。人类考试实践的结果表明，设计试题的技巧不同，测试的效果就不一样。若按人们通常采用的试题设计技巧，考试可分为两种：

（一）分离式考试

这类考试的试题设计采用的是分解设计法，即把确定欲测的考试内容，按知识的类别或能力的性质划分成众多的测试点，再根据各个测试点的要求，分别设计测试的题型及其内容，然后通过对多个知识点、能力点的测量，来实现全面检测欲测内容的目标。如美国西沙测量儿童音乐听觉能力的考试，就是将测试儿童音乐听觉能力的内容分解为：辨别音调高低、强度、时差、音色或音质、节奏和音调的记忆六个测量点，经过逐个点的测试达到对儿童音乐听觉能力的整体检测。由于这种考试采取的是分点测面的方法，各个点的测试内容一般是单因素的，所以大都选用诸如选择、填空、判断、类推、比较之类的单项试题。如上文所举西沙的儿童音乐听觉能力测试，就是通过应试者对100对音的比较结果，来判定应试者辨别音调高低的能力。

分离式考试虽有设点多、测试面广的优点，但各个点测得的质与量的和，是否能反映应试者被测方面实际具有的质与量，目前尚无足够的科学根据，仍然是有待深入探究和需在实践中进一步验证的问题。

（二）综合式考试

综合式考试的试题设计，恰与分离式考试相反，它采用的是综合设计法，即把确定欲测的考试内容相应归类后，再进行综合设计，每道试题的容量较大，至少包括两三个知识点或能力点。应试者要综合运用几方面或多方面的知识，才能正确解答一道试

题。如中国语文考试贯用的作文题，就是最典型的综合式考试试题。它对应试者的测试，不仅是对词法、句法、修辞、文字组织与表达、逻辑思维、形象思维、创造思维等多种知识和能力的检测，而且还涉及到应试者对哲学、史学、社会学、生物学、地理学等多门学科知识综合应用的能力。除作文题外，传统考试中用于测试应试者高层次知识与能力的论文题、论证题、综合应用题，以及现代考试中的综合填空题与多项选择题，也是综合式考试常用的题型。综合式考试对应试者知识和能力的测试，犹如在平面上挖井，可有效测量应试者掌握知识及所具能力的深度。但因受考试题量和测试时间的双重限制，综合式考试不易兼测应试者知识及能力的广度。为使一次考试的测试内容深度与广度兼顾，最好是同时采用两种试题设计法，既有综合性试题，又有单项试题，对应试者的测试点面结合，以保证测试结果的有效性和可靠性。

五、按测试内容的表现形式分

所谓测试内容的表现形式，系指考试内容由什么样的材料组成，或者说考试内容呈现于应试者面前的具体材料是什么，是文字、图形？还是实物？从这一角度看，考试也可分为两类：

（一）非文字考试

非文字考试所用的材料是图形、实物或口语。应试者根据主试者的要求，完成图形设计、绘画、实际操作、演示、规定的动作或口头提问等项目。非文字考试既适用于懂文字的人，也适用于不懂文字的人，用途极为广阔。属于非文字考试范畴的考试有艺术、体育院校的术科考试，外语院校的口语考试，学校系统部分学科、专业的实验、设计、操作考试，军事、科技行业的技能考试，民间各种工匠、艺人的技艺考试，以及用非文字材料编制的各种测验，如美国陆军乙种智力测验、罗夏的墨迹测验、波特

斯的迷津测验等。

（二）文字考试

文字考试也就是现今盛行的笔试，包括开卷、闭卷、问卷多种类型。文字考试要求应试者以文字形式解答主试者的文字质问。它只适用于懂文字的人，其使用范围比非文字考试窄。

六、按施测形式分

施测的形式，系指考试实施时主试者对应试者测试的具体方式。从这一角度考试可分为：

（一）个体考试

个体考试的施测，主试者是以单个考生为对象，依照先后次序逐个进行单独的测试。对每个考生来说，测试的规程、环境及试题或项目的难易度基本相同，但具体的测试内容却不完全一致，尤其是某些随机抽签作答的测试，各个应试者受试的内容差别更大。采用个体形式施测，主试者对应试者的行为能真切观察与恰当控制，除可检测应试者的知识、能力外，还能了解应试者在习惯、性格、思维模式等方面的个别差异。个体考试多用于口语、技能、技巧及实际能力的测试。耗力费时，评分误差不易控制，是个体考试的主要缺陷。

（二）团体考试

团体考试的施测，主试者是以批量考生为测试对象，以组、班或试场为测试单位，用同样的内容和要求，对所有参加同次考试的应试者同时测试。它一般是通过卷面成绩反映应试者的知识、能力水平。团体考试可弥补个体考试耗力费时的缺陷，评分误差相对小一些，但施测过程中对应试者的行为不易控制。现阶段，团体考试在社会各系统使用甚广，如在校学生的期中、学期、学年、毕业、升学考试，教师的进修、晋级、出国留学考试，企、事业部门的招工、招干、业务考试等，大都以团体形式

施测。

七、按作答形式分

迄今为止，人类所经历的考试，若从应试者作答形式的角度划分，可分为如下两类：

（一）固定应答式考试

固定应答式考试，亦即人们通常所说的客观测验和标准化考试，属客观性考试的范畴。如日本大学入学考试的第一次共同试，美国的大学入学考试（SAT）、研究生入学考试（GRE）和外国留学生英语水平考试（TOEFL）等。测试时，应试者根据考试内容所提供的信息和现成答案，经过自己的分析、理解或判断，进行填空、改错、配对或选择正确答案，以固定的形式回答主试者的质问。固定主答式考试一般采用分离式单项试题，具有题量大、内容广博的特征，对知识面的检测有独特优点，是当今世界上最为流行的一种考试。

（二）自由应答式考试

自由应答式考试属于主观性考试的范畴，传统论文式考试即属此类。这种考试的试卷中除少量单项试题外，多为论述题、论证题、证明题、综合性应用题和作文题。应试者可充分利用所学知识，根据自己的理解和认识，自由回答主试者的质问。这类测试试题的解答，论证、推理、判断性较强，应试者需有一定的文字表达和推理判断能力，并经过深思熟虑才能完成测试内容，故可深入考查应试者的文字表达、材料组织、逻辑思维能力及发散性思维能力。但由于自由应答表述的文字多，应试者需有较充裕的时间思考和组织，所以此类考试又有题量小、测试内容覆盖面窄、评分工作量大、标准不易掌握等流弊。目前，自由应答式考试正处在改造、更新、逐步向标准化考试过渡的时期。

八、按考试分数反映形式分

在现代考试中，主要有两种反映考试分数的形式，即分别以常模或规定目标为参照物反映考试的分数。从此角度划分，现行考试可归为两类：

（一）常模参考性考试

常模参考性考试亦即区分性考试，它以常模为参照点，反映每一个应试者的考试分数在同一考试总体成绩中的位置。社会上各种大规模的选拔考试和安置考试属于此类。

所谓常模，即是同类考试之间可供比较的标准，又称同类考试的标准化样本分数，亦即同类考试比较各考生分数的标准参照点。在现行的考试或测验中，常模的种类很多，最常见的有年龄常模（某一年龄的儿童，在某种考试中所得分数的平均值）；年级常模（某一年级的学生，在某种考试中所得分数的平均值）；各种规模较大的选拔性、安置性考试的常模（某一批水平基本相同的考生，在某种考试中所得分数的平均值）等。但不管上述哪种常模，它都是同类考试比较考试结果的一个参照点，用以平衡试卷、转换和反映考生的分数。

由于常模参考性考试是用某次(或某类)考试全体考生的平均分数为参照点，来反映该次考试中每个考生的分数，所以给考生评定的是相对分数。这种分数是通过与其他考生在同类（或同次）考试中所得分数的比较，反映出它在同一考试总体成绩中的位置，表明每个考生成绩的优劣，因而有利于人才的准确选拔和合理安置。有关常模参考性考试的具体内容与功用，我们将在下一章再行探讨。

（二）目标参考性考试

目标参考性考试，也称尺度参考性考试，或达标性考试。它是以既定的标准或目标作为评定考生分数的参照点。例如在学校

考试中，既定的教育目标、教学目标或课程目标，便是衡量学生阶段性学业成绩是否达到各级教育培养标准的尺度，亦即学生考试分数评定的参照点。现行社会各系统的合格考试、各级各类学校的学业成绩考试，均属目标参考性考试的范畴。

目标参考性考试与常模参考性考试，虽都是以某一标准为评分的参照物，但在性质上是有区别的。目标参考性考试是与既定目标作纵向比较的自身评价，无需依赖其它分数的比较，即可鉴别出应试者成绩的达标度。它所反映的考生分数是绝对分数，整个考试分数的分布变化以小为佳。常模参考性考试是与同类考试全体考生的平均分数作横向比较，确定自身在同一群体中的位置，所反映的考生分数是相对分数，其分数的意义有赖于与其它分数的比较，整个分数分布的离散度越大越好。前者是检验应试者的学业成绩是否达到预定目标，只论合格、不论选优；后者是通过同一应试群体的横向比较，从中甄选优秀者，只论选优，不论合格与否。因此，两种考试各有不同的适用范围。

九、按考试结果评估手段的性质分

从理论上讲，人类所进行的一切考试都是社会的客观需求，而任何一种考试的目的和内容的确立，又是主观的；苏联著名教育家赞科夫也曾针对传统考试的特点指出："学生的试卷本身是有其客观质量的，但是教师在评分时却不免掺杂许多主观因素。"再者，一切考试结果(不管是否真实反映欲测的内容)本身都是主试者与被试者两种主体交互作用的客观产物，而对考试结果的评估，又有主观与客观两种性质。以此划分现代考试的类别，又可分为：

（一）主观性考试

主观性考试，系指以阅卷者主观判断为依据评定考生分数的考试，也就是沿用至今的传统考试。由于传统考试主要采用论述、论证、综合应用、推理判断、作文之类的试题，考生自由解

答主试者的问题，同一问题的答案不拘一格，主试者不可能制出一把衡量多种答案的标准量尺，作为判定应试者考试分数的客观依据，只能凭借阅卷者个人的鉴别力区分考生答案的正误与优劣，所以考试分数中的主观成份大，同一考生的答卷给不同水平的人评阅，其分数的绝对差极为显著，如曾将一份英语考卷请145名教师评阅，所判分数的最低分为50分，最高分为94分，整个评分情况如图6-1所示。

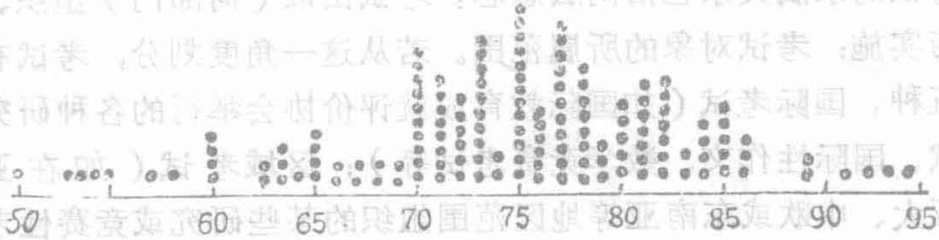


图6-1　145名教师对同一份英语考卷所评分数的分配

图6-1表明，145名教师对同一份英语考卷评分的两极差高达44分。也正是由于传统考试评分具有主观随意性大的特点，才称其为主观性考试。

（二）客观性考试

客观性考试是相对传统考试主观评分的特点而言的，即是说，客观性考试的评分不为阅卷者的主观意志所左右，它具有客观统一的评分标准和科学的误控措施：一是试卷全部采用客观性试题；二是各题的答案与分数固定，考生固定应答；三是采用光学扫描器和电子计算机阅卷评分，可控制阅卷者主观判断或运算等所造成的误差。因此，无论什么人评卷，考生的实得分数都将稳定不变。但也必须明确，客观性考试评分相对主观性考试的评分是客观的，而不等于百分之百的准确。无论采用哪种评分手段，所评分数的精确度都只能相对而言，要想没有丝毫误差，事实上不可能。这是因为：一方面，影响评分准确度的主客观因素

是多方面的，何况各种因素的可变性大，又无固有的规律可掌握。另一方面，任何一种评分手段都是在特定的历史条件下形成的，无不具有局限性。考试要以相对固定的评分手段去控制多变的误差因素，使评出的分数绝对客观，在实践中办不到。所谓评分客观，准确地说，系指所评分数能比较客观地反映被测对象实际具有的质与量。而主、客观考试，则是根据判定考试分数手段的性质区分的。

十、按考试隶属关系分

考试的隶属关系包括两层意思：考试由谁（何部门）组织、管理与实施；考试对象的所属范围。若从这一角度划分，考试有下列五种：国际考试（如国际教育成就评价协会举行的各种研究性考试、国际性作文、数学竞赛考试等）；区域考试（如在亚洲、亚太、中欧或东南亚等地区范围组织的某些研究或竞赛性考试）；国家考试（如某些国家举行的高等学校招生考试、出国留学考试等）；系统考试（如教育系统、卫生系统、人事管理系统等所举行的各种考试）；单位考试（如一所学校或一个工厂内部举行的考试）。

第二节　考试方法

人的知识、技能、体能、能力、品格、兴趣、理想、态度等外化形式的多样性，决定了考试方法的多样性。这就是说，人的知识、素质、技能与能力的外化形式是各不相同的，它需用不同的方法进行测度或甄别，才能判明被测方面量的大小与质的优劣。任何一种单一的考试方法，都不能测得人的德、学、才、识、体这一复杂总体的全部质与量。当然，人类的一切测量都是相对的，即使采取多种考试方法，从各个不同角度测量、甄别人的素质与智

能，测得的质和量与被测者实际具有的质和量，也不可能毫无差异。原因在于：人的素质与智能的外化是动态现象，不易被人捉摸，尤其是人的意识、品质的外化现象，更无稳定的特质，不少现象在人类的认识中似属“必然王国”，尚待继续研究。另一方面，测度、甄别人的素质与智能的方法，目前“也没有达到尽善尽美的地步，所以我们还不能够把知识、技能、品格、兴趣、理想、态度等，一一加以测量”①，精确测出每个人行为总体的绝对差异，还有待不断探讨、改进，但这只是方法完善和认识深化的问题，而不能因此怀疑人的素质与智能的可测性。事实上，人们通过长期的观察、试验和研究，对人的素质与智能的外化现象已有极深刻的认识，并摸索出大量符合人的素质与智能外化规律的测试方法，如问卷法、真伪法、选择法、再生法、操演法、投射法、接谈法、观察法、评定法、调查法、推定法、记录法、答辩法、联想法、社交测量法等等。上述考试方法若要一一细述，难免篇幅冗长，又重复甚多。因此，只就现行考试使用最多、彼此不重复的方法加以叙述，凡在所述方法中能予以概括者从略。

一、书面质问法

书面质问法，是通过书面文字或图表质问主试者所要测试的内容，也称之为“问卷法”。采用此法施考，先由主试者依据考试目的编制反映欲测内容的试题或项目，并按一定程序与规则拼配成试卷，再由应试者书面作答，然后对应试者的答卷进行定性、定量（或两者相结合）的统计分析，给应试者的测试结果以恰当评价。通过书面质问法测得的结果，其用途有二：作为鉴别、评定应试者素质与智能优劣的依据；或用作进行某项研究的

①孙邦正编著：《教育测验与统计》、台湾商务印书馆1986年第10版，第3页。

素材。

书面质问法，按主试者质问的目的或应试者答卷的方式，又可分为直接质问和间接质问两种：

直接质问，亦称当面质问，或谓现场答卷法，它是一种主试与被试双方面对面的书面测试。直接质问在预先确定的场所进行。施测时，主试人将问卷直接交与被试者，并当面阐明作答要求。应试者在主考人的监督下，按规定的时限，以文字、符号或图表形式现场解答主试者的书面质问。其作答方式，取决于问卷中试题的类型。若是客观性试题，应试者只能固定作答。如果问卷采用的是主观性试题，应试者便可自由作答。直接质问，虽然应试者要受时间与文字的双重限制，并在主考人的监督下作答，气氛严肃，对多数应试者充分发挥其聪明才智有一定影响，但因其施测规范，抗干扰性强，测试结果较为准确可靠。所以，它是人类考试使用极为普遍的一种方法，被社会各系统广泛使用于科技知识、技能（即属于知的要素的技能的，而不是已形成定型化行动的实技）、能力（如综合应用知识的能力、文字表达能力、聚合思维能力、发散思维能力等）、性向（如职业性向、学术性向、专业性向等）方面的测试，以获取甄别、评判应试者素质与智能个别差异的依据。

间接质问，它也是通过书面文字或图表质问主试者所要测试的内容。所不同的是：间接质问的测试目的不在于测度与甄别应试者知识、技能、能力、性向等方面的个别差异，而是通过测试获取评价应试者气质、品质、性格、志趣、态度等方面个别差异的依据，或搜集进行上述内容研究的素材；施测无固定场所，也没有时间的限制和主考人的监督，问卷以书面调查形式交与应试者；根据主试者问卷上所列质问内容，应试者按自己的理解或看法，以文字或符号形式作答，它是一种主试与被试双方背对背的书面测试；答卷上应试者可落名，也可不落名。因以上特点，人们又

将间接质问称为“自陈法”、“自我诊断法”、“自我评定法”或“问卷调查法”。这种测试方法，为英国差异心理学奠基人高尔顿所首创，后经斯丹雷赫尔的介绍，于19世纪末传入美国。间接质问对测试人的气质、品质、志趣、性格、信仰等，较其它测试方法适宜。施测过程中无需设置固定场所，也不要派专人进行控制，只要把根据测试目的编出的质问事项印成问卷，将其分发给欲测对象作答，便可从答案中观察应试者的气质、品质、性格、信仰、志趣等，获取大量的反馈信息，以供评价或研究之用。正由于间接质问的测试方法具有省时省力，以及适于测试人的品质、气质、性格、信仰、志趣、态度的优点，所以被现代考试广为采用，是现实社会进行人格测验的主要方法。

间接质问欲测的内容，是态度、志趣、信仰、道德、性格这些极为复杂而又变化无常的行为特质。应试者在作答过程中，或因对这些特质不甚了解而随意作答，或因作答与本人实际不符，或因知道该如何作答而改变自己的看法等，是常有现象。诸如此类的因素，都会影响答案的客观性。此外，人的知识与行为并不一定是线性关系，如某些道德品质低下的人，其道德知识未必比道德品质高尚的人少。基于上因，难于准确判断应试者答案的真实程度，乃是间接质问的最大缺陷。

二、口头质问法

口头质问法，是一种主试与被试直接对话式的个别测试法。它以考签、录音带、录象片、幻灯片、图片、论文等为测试的中介物。施测时，主试者根据中介物呈现的内容或情景，口头质问欲测内容，令应试者当面口头作答。然后，由主试者（或评审组）对应试者的测试结果作出评价。

若依照呈现测试内容之中介物的种类划分，口头质问法可分为四种形式：

抽签质问。这是一种使用最普遍、简便易行的口头质问形式。在施测前,主试者需将拟定的测试项目分别制成考签,测试时应试者按随机抽签的原则，抽出规定数量的考签，主试者按考签上的测试内容质问，应试者按质问的内容与要求口头回答主试者的问题。

据文质问。它以论文为测试中介，主试者按应试者事先写就的论文内容质问,高等院校和科研机关现行的博士、硕士及少数学士学位的论文答辩,即是采用此法。据文质问的规定性很强，其质问范围严格限于论文本身。以此方法施考，有着严密的程序，通常由各级学术委员会主持实施，并约请有关专家、学者参与指导。正式施测之前，主考单位必须复制论文原本，并呈送有关人员审阅，以作质问准备。施测一般分为三个步骤：首先，由应试者当众宣读或简要介绍本人的论文，再由论文评审者及其他参加测试的人员，就论文内容分别口头质问，应试者逐个口头答辩。质问完毕后，由评审组对应试者的论文与答辩进行评定，并由主持人代表评审组当场宣读书面评语。

据图质问。它以录像片、幻灯片、图片等为测试的中介物，属于情景对话或测试。这种测试方法主要用于各种语言的口语考试。在施测过程中，录像片、幻灯片、图片所展示的情景，既是主试者质问的依据，也是应试者答问的依据。面对同样的情景，主试者按自己的理解质问，应试者根据本人的理解或联想进行口头叙述，或遵照主试者的要求，有针对性地口头回答主试提出的问题。采取据图质问的方法测试口语，若同一次考试的人数较多，需有多种不同情景的画面，使每个应试者都有一定份量的新内容，以避免因质问内容重复而导致测试结果的失真。

据音质问。此种测试方法为外语听说能力考试所常用。它要求在施考前将欲测内容录于磁带，施测时通过放录音测试内容，一般包括短文、对话、提问三个组成部分。据音质问有两种方

式：一种是由主试者根据录音内容质问；另一种是录音中的直接设问。应试者亦有两种不同的作答方式：口头回答主试者的质问及录音本身的质问，或卷述录音中的短文；依据从录音中所听到的内容，以符号形式书面作答，即在试卷提供的答案中选出自己认为符合要求的正确答案，并在选定的答案上以规定符号作出标记。

口头质问法在测试人的个别差异方面，具有灵活、广泛、深入的优点，适宜于听说能力、记忆力、想象力、创造思维能力、语言技能、灵敏性以及个性特征的考查。其主要缺点是评分误差大，工作效率低。

三、交谈法

交谈法，亦称“接谈法”，它是口头质问法的变换形式，整个测试以谈话方式进行，多用于招生考试、招工考试、招干考试、特种兵招兵考试及社会和学校其它考试的面试。

与其它考试方法相比，交谈法有五大特点：其一，测试目标的范围广。交谈法既适用于不同知识、技能、技术、技艺、能力的测试，也可用来考查、甄别应试者的意识、性格、志趣、信仰、意志、理想、品德等。其二，质问内容可灵活多变。在测试过程中，围绕规定的测试目的，主试者可随意质问与欲测目的相呼应的问题，提问的深浅程度与先后次序，也可根据应试者的作答情况和交谈情绪，由主试者灵活安排。其三，应试者的文化水平不受限。不管应试者具有何种学力和文字水平，均可使用交谈法测试欲测的内容，只是交谈的具体内容，应因测试对象而异。其四，测试场所与程序可因人因地安排。它没有统一规范的施测程序，也无固定的施测场所，无论在室内室外，凡环境僻静、有利于主试与被试之间开怀舒畅交谈的地方，都可作为测试场所。其测试程序一般是根据具体对象由主试者自行确定，以便于应试

者真实回答问题为原则。其五，测试方法可长可短。采取交谈法所进行的测试，没有统一的时限要求，测试时间的长短，应视主试与被试双方交谈的实际情况决定，以完成欲测内容为限。由于交谈法有这些独具的特点，因而对应试者的素质与智能可作较为全面、深入的检测和考查。同时，此种相对自由的测试方法，又有耗费时间、质量难于控制等不可避免的弊端，而且测试结果的分析难度较大，除要具备宽厚的学科知识外，还必须有教育学、生理学及行为科学方面的基本知识和理论。再者，在分析应试者答案记录时，应切忌个人偏见的掺入。

交谈法只要使用得当，应试者能以诚实的态度回答主试者的质问，便可较为准确地诊断应试者德、学、才、识、体诸方面的优缺点，以及存在问题的原因。通过交谈法测试获取的信息，对改进教育、职业、工作上的指导有重要的参考价值，有利于在人才培养和使用的过程中防微杜渐，长善救失，因材施教，量才用人，能促进人才培养和使用效率的提高。

四、操演法

操演法带有浓厚的实践性，应试者按照一定的规程，通过现场操作、实物制作、造型设计、当场表演，分析处理实际事物或执行某项具体任务等形式，来完成主试者提出的测试项目。它在全面检测应试者的素质与智能，以及正确处理测试书本知识与检验运用所学知识分析解决实际问题能力的关系方面，是书面质问法、口头质问法的重要补充形式。操演法对测试应试者的实际技能、技术水平、艺术技巧、体能与实际工作能力等，有其它考试方法难于替代的功能。自奴隶社会以来，操演法一直为体育、音乐、美术、军事、民间技艺、技术等实践性考试所采用。在人类考试中，操演法是一种最古老而又最富现实意义的考试方法。

五、评定法

评定法，又称“评判法”和“评品法”。它以多种评定量表为工具，测量应试者各种行为的特质（即人格特点），或评定应试者体操、武术、唱歌、绘画、舞蹈等表演的技艺水平，或评价应试者各类工艺美术作品、文学作品的质量。并将测量结果的数量化、等次化。评价法早先用于美学，“现在评定法利用范围极广，从美学至工业生产，人格行动的评价乃至公路铁路评价，举凡考核评价的事，任何学校、任何机关、任何工厂，都会利用评定法。教育上应用评定法的范围颇广，尤其难捕捉之评价目标，多依赖评定法评价。”①

现行评定法所用量表甚多，不拟一一列举，下面选录几种常用的评定量表，以说明评定法的具体应用。

（一）朋路道品性测验②

“这个测验的目的，在观察你的兴趣和态度，不是智力测验，答案也没有准不准的分别。每题之前，有‘是’、‘否’、‘？’三种。若你的答案是‘是’，请将‘是’字加圈。你的答是‘否’，就将‘否’字加圈。若你以为不能决定，就将‘？’字加圈。”本测验共125题，选录前10题为例：

1．是　否　？　你不依习俗时，感觉不安适否？
2．是　否　？　你时常想入非非否？
3．是　否　？　你通常是否喜欢自寻办法，不愿求人指导？
4．是　否　？　你曾是否为避免和人招呼起见，走道街道对面去？
5．是　否　？　你能否受人批评而不感不快？

①李聪明：《教育评价的理论与方法》，台湾幼狮文化事业出版公司1973年版，第175页。

②孙邦正：《教育测验与统计》，台湾商务印书馆1986年第10版，第298页。

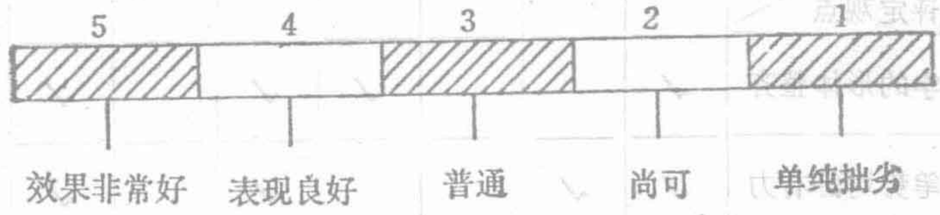
酒之會，賦詩摩崖，極林泉之樂。

予先心儀先生，己未歸自南洋，始獲親炙，遂訂交焉。尋且偕與先生相唱和，先生更稱為勁敵，固是過譽，實寓策勵於其中也。迄予作令，先生詩文投贈，難免客氣稱譽，顧猶數數隱勉以循良。先生詩宗少陵，傍及漁洋，凡所作工整清麗，如火齊走盤，才思復敏捷，日恒三數首，不經意之作。偶涉平淡，皮相者便以『雖多，亦奚以為』之句以相誚，實未知先生也。

先生處境佳，故有時對客興至，亦間作諧語。予曾於酒後談笑中，舉昔人詠文人福慧雙修之詞（一個秀才三百谷，一座書齋一連屋。窗前幾架書，屋後幾竿竹。一妻一妾三箇子，人生到此亦足樂）以戲曰：『老由秀才而貢而官，而膺二品封。租息所入不止三百石之值，宅第樓閣園林高敞清華。閣中書琳瑯充棟，園中竹茂密成林。正副配俱賢淑，子七人咸英穎。較昔人所云件件皆過，老且止足，樂且樂樂矣。』蓋潮語樂、腩同音，腩者肉爛之義，故潮人凡過敗壞糜爛而不可收拾之事，輒呼曰『腩腩』。腩者，甚之也。客聞哄堂，先生亦絕倒，雖近恢諧，盡天下讀書人得享此福，寧有幾哉？

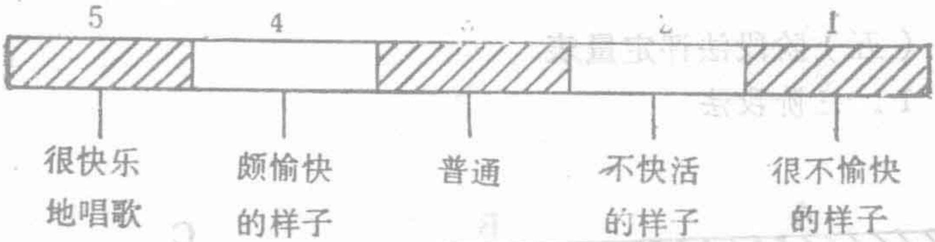
有一日已向暮，先生扶一小奚奴造予汕寓，神色惶急，一見面即大呼『不了不了』。予亟問故，先生曰：『數日前毗鄰洋客呂某致四百金丐為其母撰壽序。』予聞之摻言曰：『此可歡喜』。先生搖首曰：『喜何有？年前為張家作一篇，一日半便就，此則思索兩晝夜終不著一字，不知為何？』予曰：『老文思素捷，何遽如此？』先生曰：『張為仕宦，有事跡可供抒寫，故易。呂母乃一常媼，秖日見其紡網線、紮草團，無一言一行可紀，故難。今決謝卻，惟彼所致之金被我用去七十餘，特來向兄通融湊足璧返。』予聞之大笑曰：『此小事，以片

3、彩色的效果

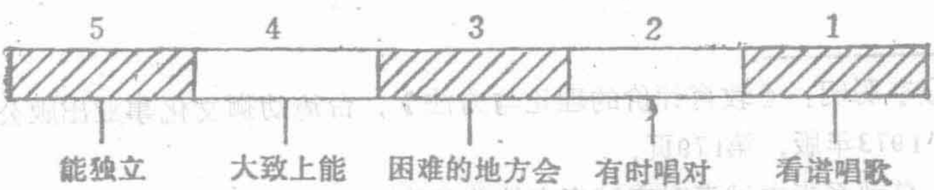


（三）歌唱方面的评定量表①

1．唱的态度——有无抱着愉快心情歌唱



2．视唱力评定



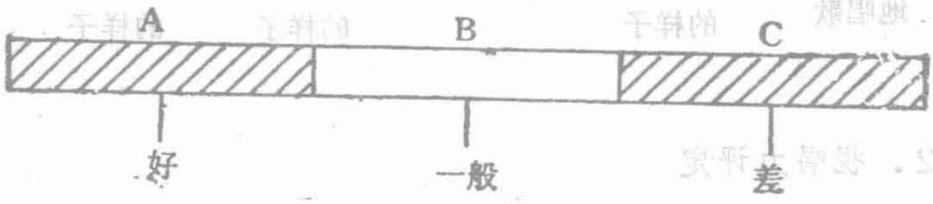
① 参见李聪明：《教育评价的理论与方法》，台湾幼狮文化事业出版公司1973年版，第176页。

（四）书法作品评定量表①

考生姓名 / 评定观点	A	B	C	D	E	F	G	H
字的形体整齐	✓			✓	✓		✓	
笔势均舒有力	✓	✓			✓		✓	✓
全部配字调和美观		✓					✓	
总　　评*	4	4	2	3	4	2	5	3

（五）阶段法评定量表

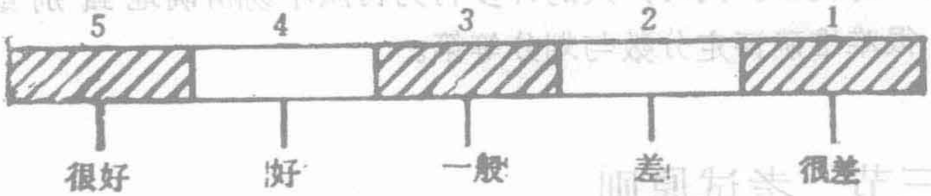
1．三阶段法



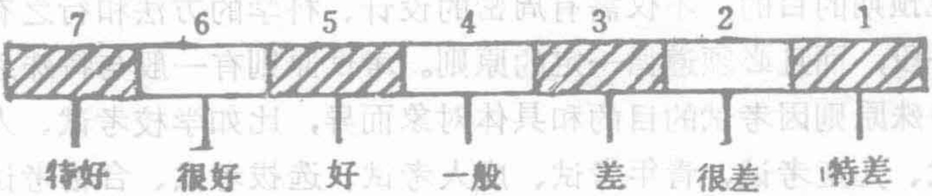
①李聪明：《教育评价的理论与方法》，台湾幼狮文化事业出版公司1973年版，第179页．

* 总评系指主试者对应试者书法作品好坏的总体评价．“4、2、3、5……”表示总体评价的分数，“5”为最高分，“2”为最低分．表中列举的三个项目，是主试者希望应试者达到的三项指标，若应试者的书法作品达到其中某一项标准，便可加1分，若三次标准全部达到，其成绩加3分，总评成绩为5分．

2．五阶段法



3．七阶段法



阶段法除以上列举的三种外，还有二阶段法、四阶段法、六阶段法、九阶段法等，但使用最普遍的是三阶段法、五阶段法和七阶段法。其中三阶段法适用于道德品质方面的评价，五阶段法适宜于评价学业成就、工艺美术作品、表演艺术等，七阶段法对划分大规模团体考试问答题的评分比例较为适宜，而在个体品质或学业成绩的评定中使用很少。一般说，阶段的数量愈少，各阶级的意义愈明确，其质与量的判断愈精确。相反，划分的段数越多，其质与量的区分越困难。因此，阶段的划分宜少不宜多，以便各阶段的意义有明显区别，每段标准易于掌握。但阶段的数量又不能太少，不然就失去了划段的意义，以至造成好坏混同、优劣难辨的局面。

对于评定法的科学性及其实际效果，目前尚无定评。一部分人认为，评定法是一种优良的测评手段，不仅适用面极广，而且结果客观可靠。多数人的看法是，对一些行为特征不稳定而又捕捉困难的评价目标，采取评定法测评要比其它测试方法有效。若评定量表制作方法科学，其评量单位标准，也能使评定的分数或

等第与评定对象的行为表现基本一致。但评定法所评结果常有偏宽偏严的现象。同时，人的许多行为特质不易精确地甄别或度量，很难准确评定分数与划分等第。

第三节 考试原则

人类的考试活动如同其它社会活动一样，要在一定条件下实现预期的目的，不仅需有周密的设计、科学的方法和行之有效的手段，而且必须遵循一定的原则。考试原则有一般与特殊之分。特殊原则因考试的目的和具体对象而异，比如学校考试、人事考试、儿童考试、青年考试、成人考试、选拔考试、合格考试，等等，都有与其对象、目的相适应的特殊原则。对此，本书有关章节已分别进行阐述。但是，凡人类考试，无论因什么目的设考，不管具体对象是什么人，也不管采取何种方式方法实施，其实质都是测度或甄别个体或团体在素质与智能方面的差异，在总体对象和目的上是一致的。因此，各种考试的特殊原则中，又无不含有一切考试活动都必须遵循的一般原则，或谓普遍性原则。其中最主要的有四条。

一、科学性原则

科学性原则是考试的首要原则，它要求人类社会各种考试的设计与实施，应符合三大规律，即社会发展规律、人的身心发展规律和考试自身的规律。否则，考试就不能产生良好的效应，甚至走向反面。其原因有三：首先，一切考试的目的及目标归根到底是由社会所提出，服务于一定社会的政治经济和文化科学，是各种考试的共同归宿。如若社会各部门设考的目的及具体考试目标的确立，脱离社会发展的实际及其对育才、选才、用才规格的

客观需求，考试就会削弱自身的社会根基，进而走向两个极端：要么降低社会人才标准，限制社会智力资源可开发的程度，成为人才智能充分发展的障碍；要么盲目提高各类人才的规格，人为扩大人才培养的投资，从而导致高才低用，造成经济和社会人力资源的浪费。反过来说，由于考试不能切合社会发展需要的实际，所选人才恰与社会意愿相悖，不能发挥一定社会所期求的功能，那么，考试也就失去了存在的价值。其次，虽然任何考试的对象都是人，但不同年龄阶段的人有不同的身心特征，不同社会部门的人有不同的知识、智能结构，不同社会和家庭环境的人对考试有不同的应试心理，即使同是一个人其不同方面的素质与智能，也有不同的外化规律与形式，考试设计与实施只有针对不同对象身心发展的特点，才能得到被试者的密切配合，使主试与被试两种主体之间的交互活动协调一致，真正用欲测的内容去测出应试被测方面客观存在的质与量，对个体素质或智能方面的差异作出较为准确的判断。第三，即使考试正确反映了社会的需求，也切合被测对象身心发展的实际，但还必须按考试自身的规律设考和施考，才有预期实现考试目的的可能。如果考试的方法不服从考试目的，内容与形式分离，随意抽掉考试的要素或更改考试的环节，同样有损考试的科学性，影响考试的效率和效益。上述三方面，既是考试科学性原则的内涵，也是考试有无生命力的根本所在，欲使考试正确行使其职能，就务必严格遵循三大规律。在中外考试史上，许多考试之所以被淘汰，原因固然很多，但从考试本身分析，违背三大规律则是导致淘汰的根本原因。

二、客观性原则

考试的客观性历来为社会所期求，也为一切考试工作者长期追求。问题很明显，一个缺乏客观性的考试，必然标准混乱、结果失真、信息有误。它所造成的损失是多方面的，可谓遗害无

穷。首先是社会受害，即因考试选才不准而使庸才得用、良才闲置。二是给育才以不良影响，甚至将后继人才引入背离社会所期望的发展方向。三是阻碍各类确富才能的人尽其智慧于社会。四是妨碍人类社会的正常竞争。那么，什么是考试的客观性呢？按罗伯特·蒙哥马利的话说，应试者考试的成绩应“取决于一个考生的能力，取决于他对所考的事物的认识，而不取决于主考人对他的答案的理解”。这是考试客观性的一方面，即判定应试者测试成绩必须客观，能准确反映考生个体被测方面知识、技能或能力的实际。除此之外，考试的客观性还包括三方面的内容：考试的依据应是客观的，应根据社会的客观需求，而不可单凭主观确定；考试的标准是客观的，就是说，度量考生素质与智能个别差异的量尺应标准、客观，不能因命题者或主试者的更替而改变量尺标准。考试的内容是客观的，所测内容必须是应试者理应掌握、且已学过、而又具有科学价值的内容，并依据可靠、正误分明。任何考试的客观性，都是以上四方面的有机统一，其中任何一种因素不客观，都会影响考试的客观性。但也应指出，考试的客观性是相对的而不是绝对的，因为考试的目的、内容、标准和评分手段，不能超越时空的限制，人们只能在社会给定的历史条件下，去认识社会的需求、选择测试的内容、确立测试标准、制订评分的原则与方法，所以，任何考试目的、内容、标准和评分手段，无不取决于两方面，一是人们认识的程度，二是社会能够提供的物质条件。也就是说，在考试目的、内容、标准及评分原则中，毫无主观因素是不可能的。强调考试的客观性，是要尽力减少考试中主观成份的比重，力求最大限度地反映被测事物的实际，更好地实现考试的社会价值。

三、公平性原则

考试既是社会育才与用才的中介，也是一切有社会实际劳动

能力的成员通往与其相适应的工作职位的检测站。因此，考试是否公平，对一个国家或民族来说，将直接影响人们的向心力，以及劳动积极性的调动和生产的效率与效益。而对应试者个人来说，则关系到权利的平等。考试的公平性原则含有三方面的要求。其一，考试的内容必须公平合理，凡参加同次考试的人，应有基本相同的答题机会和条件，或者说，应以全体考生所学而又必须掌握的材料编制考试内容，使之答题机会均等。其二，凡与施考对象条件相符合的人，都有参加考试的权利，不受种族、出身、居住地区、宗教信仰、风俗习惯等方面的限制。其三，成绩面前人人平等，凡成绩相同的考生，应享受同等待遇，而不能因人而异。

四、实践性原则

实践是检验考试质量优劣的根本尺度。一次考试的结果，究竟可靠性、有效性如何？虽可采用定性、定量分析的方法得出基本结论，但结论是否科学客观，最终还得由实践来验证。比如各级学校招生考试的质量如何，有待所选新生在高一级学校的学习实践中检验；各劳动部门招工考试的质量如何，需用被选职工在新工作中的实际成效来鉴别。总之，源于实践、服务于实践、受检于实践，是考试的普遍原则，应为人类社会各种考试共同遵循。

第七章

考试的流程、设计与实施

考试流程犹如一条完整的链条，环环相扣，节节相连，有序性是其固有的特征，必须严格按照系统的科学程序设计与实施。

考试是人类社会评量人才、选任贤能不可缺少的方法。它不仅为历代承传沿用，且为各个时代的人们倾注心思才智，继前人之基础而不断艰苦求索、革新发展，使之历久常新，日益充实，体系渐固，逐步形成了一套严密的程序和相应的原则，供人们习用遵循。如同世间其它社会现象一样，考试有其固有的内在联系和运行规律。它从确立测试目的、设计实施、直至分析处理考试结果，是一个多环节的有序流程，步骤分明，程序严密。一般说，考试的目的、内容、方法、手段等，是根据实际需要及所具条件而定，但考试的程序、步骤却是相对稳定的，不能随意颠倒序次、调换步骤，必须严格按照系统的程序设计与实施。

第一节　考试流程

人类社会现行的各种考试，是因社会育才、选才、用才的不同需求而设立的，各有其特殊的目的和内容，设计与实施的某些

具体步骤，也不完全一致。同时，由于社会对测度、甄别人才素质与智能个别差异的要求，是随社会生产的发展和分工的变化而不断提高的，为使考试能动地适应变化发展的社会需求，人们仍在为实现考试程序的科学化而竭智尽力，考试流程中的若干细小环节，始终处于改革之中。基于上述原因，目前不同国家、地区、乃至社会各行业间现行考试的流程，并无举世公认的固定模式。若弃小异而取其大同，人类社会任何一种规范化考试的实现，则必须历经这样一个有序的流程：确立考试目的──→设计测试内容──→组织施测──→评定考试成绩──→分析处理考试结果。为说明考试流程之共性结构与关联，拟将按考试目的归并的六类考试（安置考试、合格考试、成就考试、选拔考试、诊断考试、研究性考试）所常用的两种流程模式（目标参考考试流程模式、常模参考考试流程模式），分别示以流程图。

一、目标参考考试流程模式

目标参考考试，由于它是以课程、教学、教育目标或其它某种既定的标准，作为评定和解释考试分数的参照点，来检测应试者的实际水平达到何种程度，无需在考试过程中再制定检测标准。

目标参考考试的最大特点，就在于考试目标是如实反映考试外部某种既定目标，而不是根据考试自身目的的需要由设考者重新制定考试标准。

这种既定目标，在目标参考考试中有两种用途：一是作为评定和解释考试分数的依据；二是作为衡量应试者学业水平达标度的标准。

因此，其流程有别于常模参考考试。目标参考考试又因是否采用题库试题而分为两种流程模式。不采用题库试题，需有编、审试题和预备测验两个步骤；采用题库试题，可减少这两个步骤，如图7－1和图7－2所示。

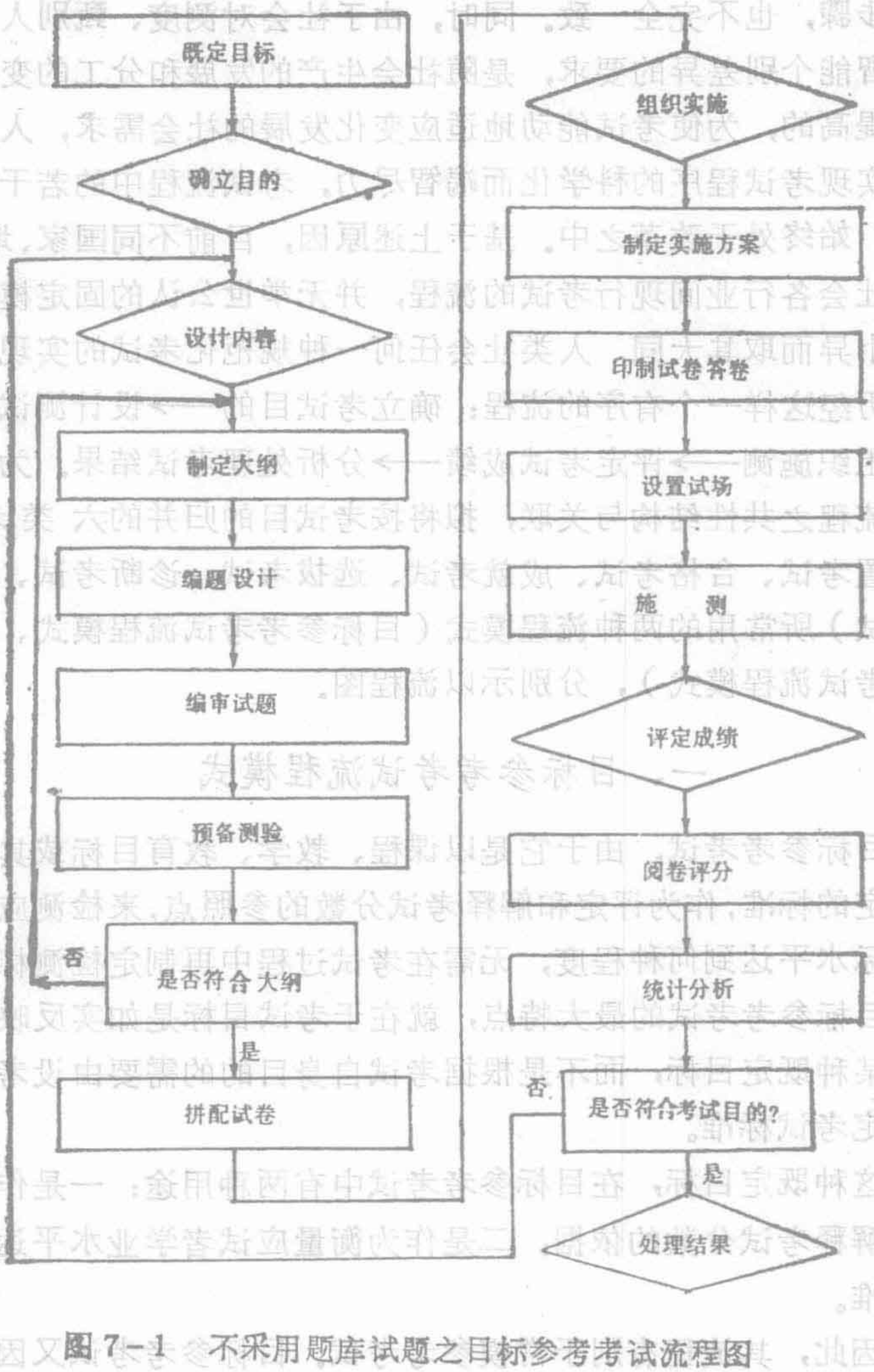


图7－1 不采用题库试题之目标参考考试流程图

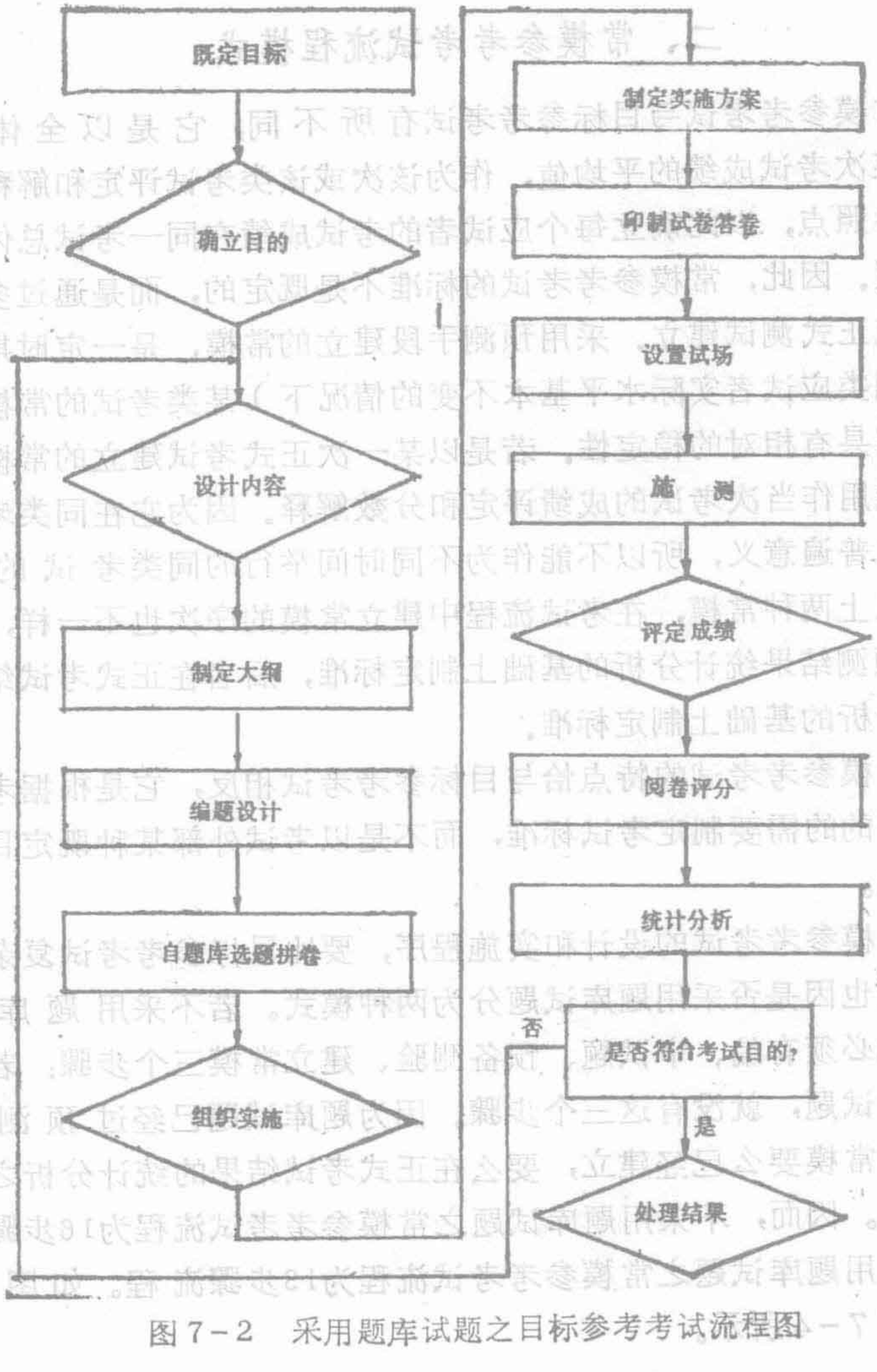


图7－2　采用题库试题之目标参考考试流程图

二、常模参考考试流程模式

常模参考考试与目标参考考试有所不同，它是以全体应试者某次考试成绩的平均值，作为该次或该类考试评定和解释分数的参照点，以此确立每个应试者的考试成绩在同一考试总体中的位置。因此，常模参考考试的标准不是既定的，而是通过多次预测或正式测试建立。采用预测手段建立的常模，是一定时期内（在同类应试者实际水平基本不变的情况下）某类考试的常模，其标准具有相对的稳定性。若是以某一次正式考试建立的常模，它只能用作当次考试的成绩评定和分数解释。因为它在同类考试中没有普遍意义，所以不能作为不同时间举行的同类考试的常模。以上两种常模，在考试流程中建立常模的序次也不一样，前者在预测结果统计分析的基础上制定标准，后者在正式考试结果统计分析的基础上制定标准。

常模参考考试的特点恰与目标参考考试相反，它是根据考试自身目的的需要制定考试标准，而不是以考试外部某种既定目标为标准。

常模参考考试的设计和实施程序，要比目标参考考试复杂。其流程也因是否采用题库试题分为两种模式。若不采用题库试题，就必须有编、审试题、预备测验、建立常模三个步骤；若采用题库试题，就没有这三个步骤。因为题库试题已经过预测检验，其常模要么已经建立，要么在正式考试结果的统计分析之后再建立。因而，不采用题库试题之常模参考考试流程为16步骤流程；采用题库试题之常模参考考试流程为13步骤流程。如图7－3和图7－4所示。

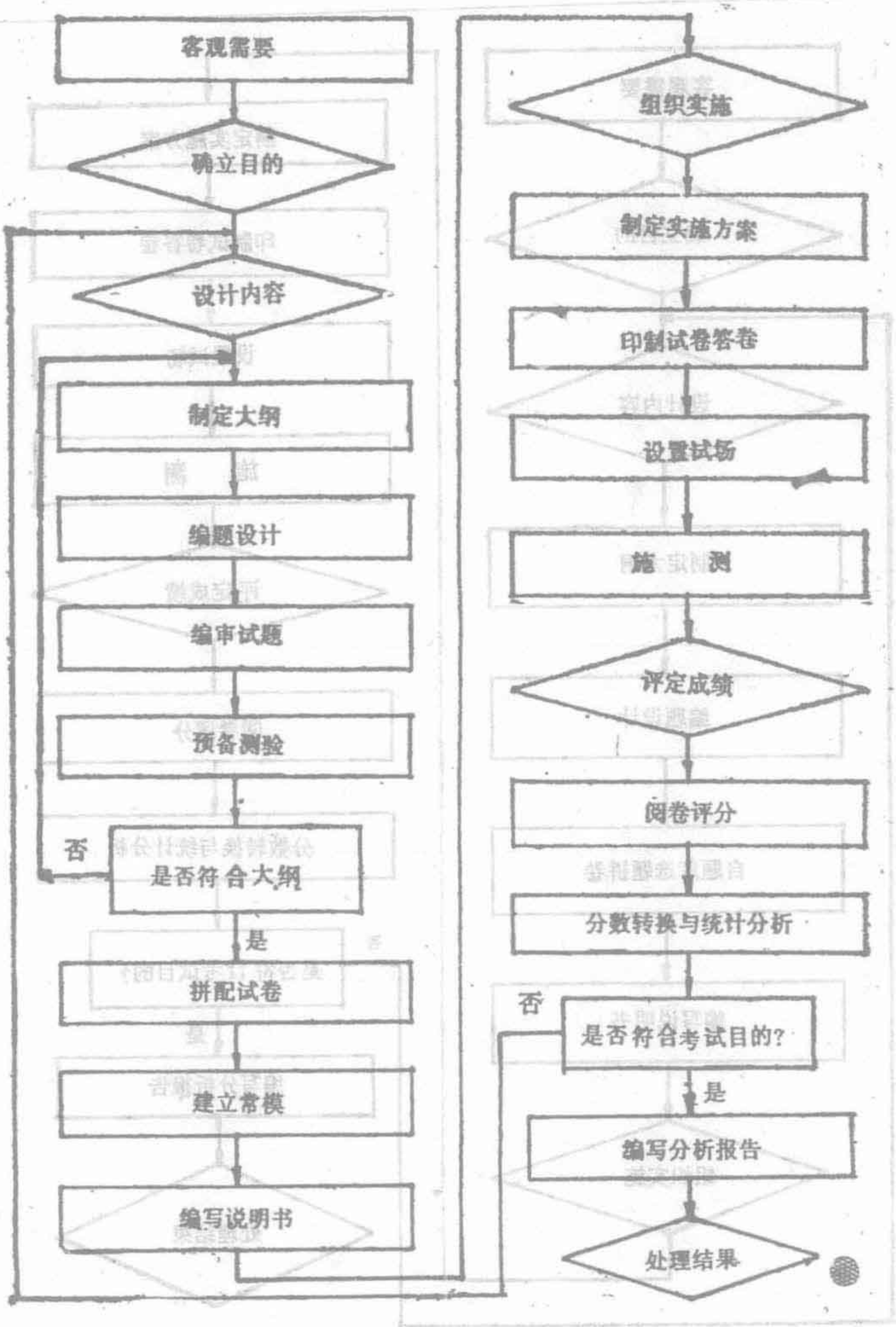


图7-3　不采用题库试题之常模参考考试流程图

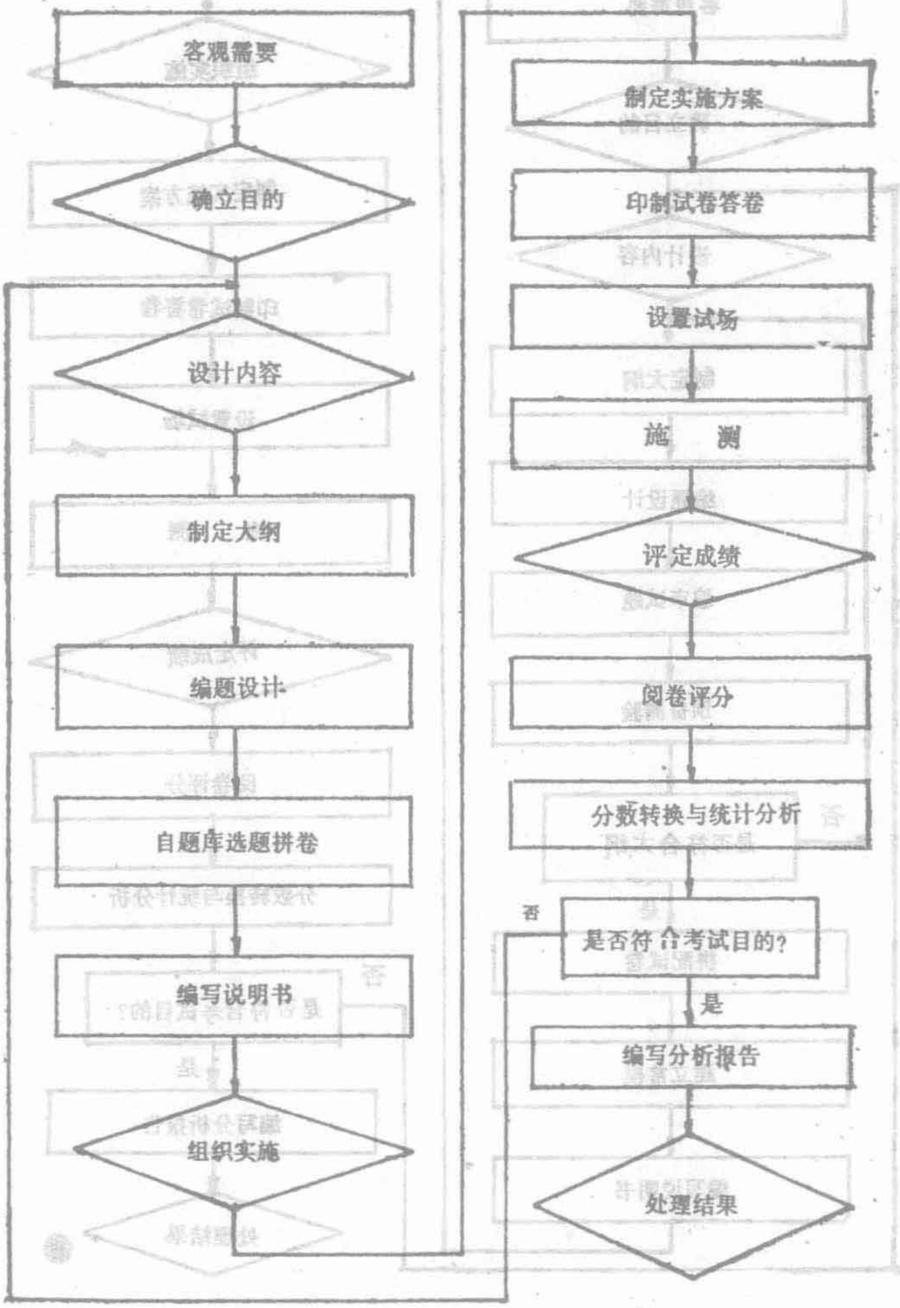


图 7－4　采用题库试题之常模参考考试流程图

第二节　考试设计

所谓考试设计，实质是要解决实现考试目的的法则与措施，它是一项科学性、技术性很强的工作，并不是简单套用上述程序设计，就能使各种考试圆满实现预期目的。前文列举的几种考试流程模式，仅仅是展示了两类考试流程的基本轮廓，或谓主要步骤。实际上，考试的流程远比这复杂。就考试设计方面讲，流程图中所列步骤，它只反映出考试的一般程序。这些大的步骤还包括若干细小环节。而每一环节的具体内容、技术要求及制作方法，又因考试目的、性质的区别而不尽相同。下面就考试设计中一些带有共性特征的主要问题，按照上图所列考试设计步骤的序次，分别进行讨论。

一、确立目的

考试是人类社会一种有意识地测度、甄别人的德、学、才、识、体个别差异的活动。任何一种考试，无论它采用何种内容、法则，也不管在什么样的场所、对象中进行，都要受到某种意识的支配，也无不针对一定目的。例如，各种交通工具驾驶执照的考试，目的在于检验应试者掌握驾驶知识、技能的准确与熟练程度，即驾驶技术是否达到规定的合格标准；工厂的招工考试，目的是要从应试者中挑选符合某工种素质与智能要求的新工人；学校在新生中进行的学前考试(或称安置考试)，目的是了解入学新生实际水平的个别差异，以恰当安排教学班、组；教师在教学过程中举行的诊断考试，是以了解教学中存在的问题、困难及改进教学为目的的。从另一角度讲，在人类的考试史上，凡是在某种意识支配下进行的考试，其内容、方法、手段等，都是为实现既

定的考试目的服务的。或者说，考试的对象、内容、方式、方法与手段，只能根据既定考试目的的客观需要抉择。各种考试的设计，之所以把确立考试目的作为首要步骤，正是因为没有明确的考试目的设计就无所适从，只有确定了考试目的，设计有了明确的指导思想和依据，才不致使考试出现形式与内容不符、方法与目的脱节，愿望与效应相悖的不良后果。

在明确三大问题的基础上，作出四项决定，是确立考试目的这一步骤的基本内容和主要任务。

（一）明确为什么考，决定考试类型

明确为什么考，就是弄清设考的依据，即举行考试的客观需要何在？这是确立考试目的应首先明确的问题。因为考试目的取决于客观需要，考试类型又从属于考试目的，假若是为获取颁发某级教育毕业证书这一客观需要而设置的考试，显然，检测应试者的学业水平是否达到某级教育规定的培养目标，便是该考试的目的，那么，考试的类型就应选用合格考试。如果设考的目的在于分配某批新工人的职位，那么，其考试类型就只能采用安置考试。若是以其它目的设考，则应在诸如选拔考试、成就考试、诊断考试、研究性考试等类型中另作选择，以符合既定考试目的要求为原则。以上说明，弄清为什么考是决定考试目的及其施考类型的前提。

（二）明确考什么人，决定应考条件

考试是主试与被试两种主体彼此交互的活动。因此，当解决了为什么考的问题，并确定了考试类型后，就应着手生源情况的分析，明确施考的对象，即准备在什么范围、对什么人进行考试？有无足够的生源？设想中的考试对象的现实状况如何？对此必须细致分析研究，做到心中有数。然后根据生源的实际情况及考试目的的要求，对应考者的年龄、学历、政治和身体条件等，作出具体规定。若是学校内部考试，便只需根据教学的要求规定

施考的年级或班、组即可。但对施考对象学习的实际状况和身心特点等，也务必中肯分析。否则，就不能使考试有的放矢。

（三）明确考什么，决定考试内容

这也是考试设计必须首先明确的问题。明确考什么，也就是明确所设考试欲测的目标，比如，是考知识、技能、学术性向、体能？还是考能力或品质？以及考哪些方面的知识、能力或技能？在此基础上，再决定考试的内容与范围。如测试不同方面的知识、技能与能力的具体内容是什么？用什么东西测试以上欲测内容的“点”？用哪些内容去考查与“点”相关的“面”等，都要有明确的规定，为编制试题初步定调。

（四）决定考试方式与方法

明确以上三大问题和有关内容，是确立考试目的这一步骤的核心所在，它关系到考试是否符合客观需要、能否适应测试对象的实际、可否取得预期效果等实质问题。同时，也只有明确了为什么考、考什么人、考什么之后，才能决定考试的具体方式与方法。比如，在考试方式方面，是采用笔试还是采用面试，或笔试、面试并用？在测试方法方面，是使用书面质问法、口头质问法、交谈法，还是使用操演法或评定法？由于考试的方式、方法牵涉到试题或测试项目编制的方式与方法，以及试题或测试项目的类别与内容，直接影响测试效果。因而，对考试方式、方法的决定应持慎重态度，使之切合考试目的的要求。

二、制定大纲

制定考试大纲，不是人类所有考试设计必经的步骤，而是现代考试对传统考试设计程序的创新和发展。即使在现代的考试中，也不是凡考试设计都需制定大纲，象学校考试系统的学前考试、平时测验、诊断考试、期中考试、学位论文答辩，社会考试系统的安置考试、职业性向考试、专业技术考试等的设计，一般

没有制定统一考试大纲的步骤。只有那些规模较大的、定期循环举行的、具有相对稳定性的、或作周期性比较的考试，才在考试设计时制定统一的考试大纲，如现实社会用于人才选拔、教育评估、文化普查、毕业会考、学年考试、学科考试等方面的常模参考考试和目标参考考试，即属此类。

考试大纲，是指导考试设计、实施、备考和应考的总纲。其内容包括：考试的性质、类型、方式与方法；测试的对象、内容及范围；试卷的长度、结构、难易度、分数比例、题型及其比重；施测时限、答题要求、评分手段、计分方法等。有的考试大纲还附有试卷样本，以体现大纲的内容与要求，便于具体执行者或参考者准确理解大纲。

考试大纲的制定，是一项严肃慎重而又极为细致的工作，一般是在考试专家和各学科专家的指导下，由专门班子根据考试目的的要求（若是为教育系统设置的考试，尚需参照有关教学大纲的要求），经过反复调查研究、分析论证后，形成文字决定，再由专业考试机构或教育主管部门审定，并通过报刊或以小册子的形式提前公诸于众，作为考试设计人员编制试卷的依据，指导考试的实施，并供学校组织教学和进行应考训练参考。考试大纲一旦确定并公布实施后，就具有法规的性质，应成为主试与被试双方共同遵守的总章程，务必照章执行，不得随意改动。若因客观条件变化，原定大纲确已不适应社会发展的要求，需要修改或重新制定，也要严格按照上述程序和原则进行。在新大纲尚未公布实施之前，仍以原定大纲为据。

现代一些规模较大、影响强烈的考试(尤其是标准化考试)，之所以把制定考试大纲列为考试设计的重要步骤，是因为考试大纲对实现考试科学化有突出的功用。这主要体现在三方面：其一，保证考试质量，使考试产生正常的社会效用，赢得良好的社会信誉，维护考试的权威性；其二，稳定考试标准，能较好防止

考试分数的升值或贬值，有利于同类考试质量的分析比较和工作的改进；其三，科学的学科考试大纲，能弥补教学大纲的某些缺陷，给教学导以正确方向，强化师生双方追求教学、教育目标的意识，促使教、考、学三方目标一致，有利于考试教学功能的正确发挥。

三、设计蓝图

设计试卷蓝图，又称编题设计，实际就是根据考试大纲的要求，将欲测试的内容和目标制订出一份双向细目表（双向特规表、试卷蓝图、测验蓝图），作为指导编题、制卷的具体计划。如果说考试大纲是考试这一系统工程施工的总体方案，那么，编题设计便是绘制施工的图纸，它是考试大纲中有关测试内容、范围、题量、题型、时限等原则规定的具体化，或谓图表化。将测试内容的类别，不同内容在试卷中所占的份量，各种内容采用的题型，各类试题应占的比重，认知能力的层次及比例，分数比例和时间分配等，分门别类地定出量化指标，并以图表形式作出详尽明确的反映。知识完整，目标层次合理，各种内容的比例适当，试卷结构严密，难易适度，时间分配科学，是编题设计的基本原则。

编题设计的目的，在于克服命题和判卷的主观随意性，控制试题和试卷的质量，保证考试的可靠性与有效性。编题设计要绘制出实现此种目的的施工图纸，首先应决定考试的内容、范围和目标层次，明确考试欲测的知识和能力。所谓知识，若从“完整知识”的概念讲，应包括基本术语和概念，一般活动的事实和科学事实，基本的科学规律，基本的科学理论，关于活动方面的知识，关于评价客观现实的知识等，而对某一具体学科来说，“指的是某一学科的各个课题，所课能力指的是通过教学在认知行为上要达到的目标。”①也就是人们所称的教育目标。现行考试一

般是根据美国心理学家布鲁姆提出的教育目标分类学说拟订编题计划。布鲁姆认为，教育目标可以认知活动作为分类的基础，他曾在《教育目标分类学》一书中阐述了这一观点，并将学习的认知活动分为六个不同级别的层次（或六种不同认知能力的教育目标）。其分布形势如图 7－5 所示：



图 7－5　教育目标（认知能力）层次分布示意图

图 7－5 各层次（类别）的基本含义是：

（1）　知识。也叫记忆或认知。是指记忆、重认与辨识事物的能力。如认识具体事实，知道有关定理、原则、术语、概念、方法、过程，对字、词、符号、人物的重认，知道事物的特征、结构与类别等。

（2）　理解。系指掌握所学材料之内容、概念、意义、含义的能力。它包括转译、解释、归纳、推想等方面的内容。如将有关数学、化学的文字表述转译为符号公式，将古汉语译为现代汉语，中、外文互译，解释有关概念、原理、法则、公式、图表、

① 国家教育委员会学生管理局主编：《标准化考试简介》，高等教育出版社1985年版，第14页。

不同文体的方法与程式，归纳文章章节大意，分辨相关或相近知识内容的异同，推想某项事物发展、延伸的趋向及其后果等。

（3）　应用。是指将所学知识用于某种实际的能力。如利用已知原理、方法、定律等解答试题，或根据给定的条件进行工程、工艺设计，绘制出新的图表，或将知识要素性质的技能用于某种实技等。

（4）　分析。是指将所学材料的内容分解为不同的组成部分或因素，以了解它们的内在结构及其功能的能力。分析包括三类：要素分析（因素分析）。它要求确认内容中的组合成份，或判明内容的组成要素，如指出某事物发展演变的成因，澄清某种定理的组合要素或假设；关系分析（相关分析或关联分析）。即找出各组成部分之间的内在联系及相互作用；系统分析（原理分析）。确定事物各要素组合、排列的原理、原则及方法，包括事物外化的与稳含的结构模式及其功能。

（5）　综合。指的是将各部分进行组合排列，并形式某种新的模式、结构或整体的能力。如综合运用所学知识及已知条件制订一项新的计划、方案，或设计一个新的程序系统，根据要求综合解答问题，通过综合运用所学知识结合给定的材料写出结构新颖的文章等。

（6）　评价。评价能力（评鉴能力或判断能力），系指按照某种既定标准或根据一定的目的判断事物、结论、法则、观点、作品的价值，判定理论正误，评价文字材料的层次结构、逻辑一致性等。评价能力，是认知能力的最高目标层次。

布鲁姆关于认知目标的六个层次是依序排列的，从知识到评价逐级递进，前一层次的认知能力是后一层次认知能力的基础，后一层次的认知能力是前一层次认知能力的发展和提高，达到了前一层次的认知能力，才有可能上升到后一层次的认知能力。因之，若认知能力达到了后一层次的目标，也就表明前一层次认知

能力的目标已经达到。

编题设计的第一步，一般是确定考试的内容与目标，即整份试卷应达到的功能，继而决定每一种内容与目标的题量、题型，再决定每道试题的测试内容和考查目标，最后根据确定的内容和目标编订双向细目表，通过纵、横两个坐标，分别标明内容和认知能力两种向度，并将不同内容和能力目标细化为具体指标，在表格中分类分项列出，以作编题和制卷的依据。

双向细目表的制订并无固定不变的模式，设计者可根据考试的要求与需要，采取不同形式标明欲测内容和能力目标的两个向度。

另外，不同的考试有不同的目的，对测试内容和能力目标各有特殊要求，不是所有考试都要达到布鲁姆提出的六级认知能力的目标，故不能把六级认知目标当作刻板的公式机械搬用。不同考试的编题设计，在欲测知识、能力方面的具体内容与目标，以及不同内容、目标在试卷总体中所占的比例，都要根据考试目的和对象的要求确定，应特点各具，如选录的表 7－1 力学双向细目表、表 7－2 高中语文科双向细目表、表 7－3 高中数学科双向细目表的内容、目标及设计形式都不尽相同。

从以下三份双向细目表可看出，它们的设计形式各不一样：力学双向细目表和高中数学科双向细目表，是以横坐标反映欲考能力目标的向度，以纵坐标反映欲测内容的向度；高中语文科双向细目表刚好相反，它是以纵坐标反映能力向度，用横坐标反映内容向度。三种考试欲测的内容和认知能力层次，也不尽相同。但它们测试的重点及其比重的分配上，却基本一致，都突出了理解、分析、应用、综合几个认知能力层次。如力学考试的理解、分析、综合三个层次占65%，而记忆只占35%；高中语文考

表7－1　　力学双向细目表①

能力目标 百分比（题数） 内容	记忆	理解	应用	综合
力概念的演变 5%（2题）	100（2）			
力的类别 10%（4题）	50（2）	50（2）		
二维力 20%（8题）	25（2）		75（6）	
三维力 25%（10题）	10（1）		90（9）	
物质的相互作用 40%（16题）	40 （6或7）	50（8）		10 （1或2）

试理解、分析、应用、综合四个层次占70%，其它占30%。同时，编题设计所决定的一个考试的欲测内容和能力目标，一般是成正态分布，六个认知层次在试卷总体中的比重，应是最高、最底两个层次份量小，中间层次的份量大，图7－6是五种考试不同认知层次的比重。

①参见香港肖炳基：《教育测量和标准化考试讲义》第2讲，第4页。

表 7－2　　　　高中语文科双向细目表①

题数　内容 能力	史论〔文言〕	哀祭〔文言〕	游记〔文言〕	议论〔白话〕	文评〔白话〕	哲理〔白话〕	小品〔白话〕	词曲〔文言〕	合计	百分比
意旨探讨	1 ①		1 ⑮	1 ⑳	1 ㉑	1 ㉛		1 ㊴	6	15
词意理解	1 ②		2 ⑫⑬			1 ㉗	1 ㉝	1 ㊲	6	15
内容分析	2 ③④	3 ⑥ ⑦⑧	1 ⑪	2 ⑰⑱	2 ㉓㉔	3 ㉙ ㉖㉚	1 ㉞		14	35
综合推理	1 ⑤	1 ④		2 ⑯⑲	2 ㉒㉕		1 ㉟	1 ㊳	8	20
文体鉴别		1 ⑩						1 ㊵	2	5
修辞辨认			1 ⑭			2 ㉘㉜		1 ㊱	4	10
合计	5	5	5	5	5	7	3	5	40	100

①参见香港肖炳基：《教育测量与标准化考试讲义》第2讲，第6页。

表7－3　　高中教学科双向细目表①

能力 份量 内容	概念掌握	证明	空间	演绎推理	归纳推理	综合分析	实际应用	总计%
1　数、代数式	1					2	2	5
2　解方程	1					2	2	5
3　指数、对数	1	2		1	1	1	2	8
4　三角形、等边形	1	1	1	1	1	1	2	8
5　圆	1	1	1	1	1	1	2	8
6　三角函数	2	2		2	2	2	6	16
7　解析几何	1	2	1	2	2	2	6	16
8　立体几何	1		2			1	1	5
9　极坐标参数方程	1	1				1	2	5
10　数列、极限	1					1	2	4
11　排列、组合	1					1	2	4
12　集合	1					1	2	4
13　二项式定理	1					1	2	4
14　微积分	1	1				2	2	8
总　计　%	15	10	5	8	8	19	35	100

① 参见李德义：《标准化考试简介》第2页。

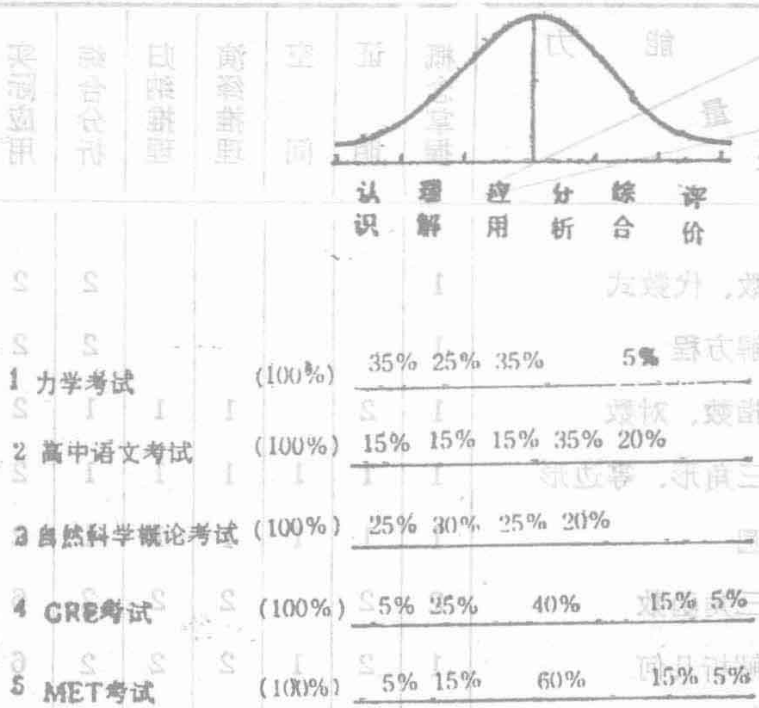


图7－6 五种考试知识和能力的比例分布

前文所述双向细目表之制作方法，主要用于客观性试题试卷的编题设计。非客观性试题试卷双向细目表的制订，虽设计程序类似客观性试题试卷的编制，但由于论文题一般综合性强、涉及面宽，一道试题所含知识和能力的测试内容及分数比例，不能量化为一个指标，要分为多个采分点。所以，在制订非客观性试题试卷的双向细目表时，还应根据每道试题欲测的知识内容和认知层次，确定各采分点的分数比例，并以二维蓝图标明。现举香港1974年历史科会考一道论文题为例，借以说明论文题分数在二维蓝图中的分布。

试题原文："蒙古大帝国，幅员广大，对于交通设施，异常重视，试于附图填上元代对外海陆交通之重要路线，并加以说明。又元代对外海陆交通于东西交流有何贡献，试分析之。"

这道论文题共24分，其各采分点分数分布及比例如图7－7所示。

双向细目表的确定，意味着编题设计的结束，即可以双向细目表为施工蓝图，依次进行试题的编写、审定，以及试卷的拼配和复本的编制。

四、编制、审定试题

编制、审定试题（命题）是考试设计的关键环节。这一步骤与考试的成败密切相关。因为无论是传统考试还是现代考试，其本质都是根据一定的目的，用一定的内容，选择一定的方式方法，去测度、甄别应试者德、识、才、学、体诸方面的个别差异。一份试卷，或一组测试项目，便是测量应试者某种行为特质的一把量尺。所以，编制考试试题，即是制作测试量尺，审定试题，亦即检查所制量尺符合标准与否。考试的长期实践表明，考试结果的真实性，主要取决于测试量尺的准确性，测试量尺的精确度，又决定于试题编制和审定的方法和技巧。这就要求试题的编制与审定，应有科学的原则和程序。

（一）试题编制的基本原则

随着人类考试实践的发展，试题的类型不断翻新，现行考试所用题目的形式，可细分为数10种。若按判分手段的性质及应试者作答的方式，也可将现行多种形式的试题归为两大类：客观性试题和非客观性试题（以判分手段的性质分），也称固定应答型试题和自由应答型试题（以作答方式分）。两类试题异同相兼、利弊各殊，功能及适应范围亦有不同。两者异同之大要可通过表7－4作粗略比较：

表7－4　　两类试题特点比较

内容＼题类	客观性试题	非客观性试题
两类试题特点	1．题量大，内容单一明了，可测面宽，易测"面"不易测"点"。	1．题量小，内容综合程度高，可测面窄，易测"点"不易测"面"。
	2．应试者固定回答主试的质问。	2．应试者自由回答主试的质问。
	3．命题难度大，技术、专业性强，耗力费时，评分客观，易采用电脑技术和量化工具。	3．命题较简便，损时损力；评分要求高，不易采用电脑技术和量化工具，工效低，主观随意性大。
	4．应试者主要精力用于试题的阅读、辨认、分析和推断。	4．应试者主要精力用于答题的思考、组织与表达。
	5．重在检测应试者知识、智能的结构。	5．重在检测应试者掌握知识的数量与质量，技能的准确与熟练。
	6．不易检测发散、组织、评价、独创、综合论证、文字表达等高层次认知能力。	6．能检测高层次认知能力，有利于特殊才能的发现和个性的培养。
	7．试题标准稳定，分数等值，能加减，也能互比。	7．试题标准上下波动，分数不等值矛盾突出，缺乏可加可比性。
	8．测试内容宽，目标分散，信度较高、效度较低。	8．测试内容窄，目标集中，效度较高、信度较低。
	9．误差易控、结果较准确，效应性较好。	9．误差难控，结果不甚真实，效应性较差。
	10．适用范围受到一定限度。	10．可供各种考试使用。

客观性试题与非客观性试题，以及它们各自所包括的多种多样的试题，尽管特点各具，差异是多方面的，试题编制的具体要领与技巧也各有特殊。但是，无论是属于何种类型的试题，其本质、属性、作用对象和编制步骤是不变的。在本质上，凡试题都是测试内容的载体，或主试与被试间的媒介；在属性上，凡试题都具有主、客观双重性质；在对象上，凡试题的作用对象都是人的素质与智能；在程序上，凡试题编审都要历经明确要求、备材、编写、审定四个步骤。因此，在试题的编制过程中，下述基本原则，应为一切编题人员所遵循：

（1）试题的形式必须符合考试的目的。试题内容应是能测出欲测量的知识和能力的内容。

（2）试题内容的所涉范围及不同内容的比重，应以编题计划为准，不得随意扩大范围、增减份量。

（3）在同一试卷中，同类试题的编写格式应规格统一，不能因试题的格式有异而导致考生审题的误解。

（4）试题要有针对性，切合应试者的生理和心理特征，试题难度必须与测试对象的水平相符。

（5）试题必须具有实际意义，不能违背科学，或夹杂不健康的成份，以致失去教育性、产生不良影响。

（6）试题内容的表述，必须用词恰当、文字简练、表意确切明了，与解题无关的字词不留，与解题相关的字词不漏。

（7）所编试题，不能存在与教材、公开印行的复习资料、教参资料中完全相同的文字表述，更不允许原封不动地照搬有关书刊、材料中的试题。

（8）试题的作答要求和指导语，要言简意明，不致使应试者费解或误解。凡需以特殊方式解答的题，应在指导语中明确提出要求。

（9）所编试题应各自独立，各试题之间不得互有关联，试

题和答案的文字表述不能有重复现象，也不能给本题或同卷其它试题提供正确作答的线索。

（10）试题内容的代表性要强，必须重点突出、兼顾全面，既测知识又测能力，不可偏废。

（11）若不是为题库编制试题，而仅是编制某一次考试需用的试题，所编试题的数量至少是最终需要量的2～3倍，以供审定时筛选和用于同次考试的复本制作。

（12）试题应便于施测、作答，阅卷评分省时省力，抗干扰性强。

（二）客观性试题的常用题型及编制要领

客观性试题的突出特征有三点：一是大多数试题的答案为设题者事先所提供；二是考试结果的评价客观准确，不论用何种方式阅卷评分，均不受评阅者主观意识的干扰，应试者实得分数不变；三是固定应答，试题既提供测试内容，同时又提供答案，应试者根据自己对主试者质问的理解、分析或推断，从主试者事先拟定的考试答案中，选出自己认为正确或符合质问要求的答案。这类试题的编制材料可用文字、公式，也可用图形、图表，它适用于多种知识、技能、能力的测试。客观性试题的变式很多，常用的题型有以下数种：

1．正误题。

正误题，也叫是非题、正负题或辨别题。试题通常是提出一个论点、一个概念、一个公式或一个史实，要求应试者判断正误、辨别真伪、分辨是非。（√表示对，×表示错，根据判断在括号中写上对或错的符号）

例1：菊花生长在水中。　　（　　）

例2：凡是掌握了语法和修辞知识的人，都能写出好文章。　（　　）

例3：西厢记是施耐庵所作。　　（　　）

2．填空题。

填空题，是由设题者提出一个意义不完整的陈述，其中略掉一个字、一个词，或几个字、几个词，并将相应位置空出，要求应试者在空白处填上符合题意的字或词，故又称此种试题为填字或填充题。

例1：短句有两种：一为________；一为________。

例2：景德镇以产________著名。

3．删字题。

主要用于文法修辞方面的测试，设题者一般是提出一个有语病的陈述，让应试者删去多余或不当的字词。

例1：王明正在读看报。

例2：李勇喜欢坐块椅子。

4．简答题。

简答题是填充题的一种变式，即由设题者提出一个疑问，令应试者在空白处填写答案。

例1：等角三角形的每一内角是多少度？　（　　）

例2：以、于、与、为四字属于哪一种词类？　（　　）

5．补充题。

设题者提出一个不完整的句式，要求应试者在空白处填上未完部分或结论部分。

例1：发明钻木取火的是___。

例2：以三直线围成的图形是一个____。

6．对偶题。

设题者在同一试题中提出多个问项和多个答案选项（答项的选项与问题的数量可对等，也可大于问项），应试者需选出自认为与各问项相对应的答案。根据设题者作答要求，应试者或用线条将问项与对应的答案连接，或把与问项对应的答案符号填入作答处。这种试题最适合名称、公式、日期、事件、地点、史实、定义、规则、符号等方面的记忆、联想、辨认能力的检测。若设计

科学，对偶题可测试不同层次的认知能力，并有节省试卷版面的优点。

例1：请用等号连接下列各几何公式：

①圆的面积	直径的三方的$\frac{3}{4}$×3.1416
②椭圆的面积	长短二轴的乘积×0.7854
③正多角形的面积	$\frac{1}{2}$（底边×高）
④平行四边形的面积	$\frac{1}{2}$（各边长之和×边心距）
⑤梯形的面积	直径×3.1416
⑥三角形的面积	半径的平方×3.1416
⑦圆周	直径×圆周
⑧圆的直径	$\frac{1}{3}$（底的面积×高）
⑨圆柱的侧面积	高×$\frac{1}{2}$（平行二边之和）
⑩圆锥或锥的侧面积	圆周÷3.1416
⑪球面积	圆底的周×高
⑫圆柱的体积	$\frac{1}{2}$（圆底的周×斜高）
⑬球体积	底面积×高
⑭圆锥的体积	底×高

例2：把右边各事件的字母代号填入符合左边时间或人名的括号内。

（　　）18世纪初	A．发现美洲
（　　）18世纪末	B．美国首位太空人
（　　）格　兰	C．美国国父
（　　）哥伦布	D．首届美洲大陆国会
（　　）华盛顿	E．美国宪法草拟成功
	F．发明电话

7．多选题。

多项选择题在客观性考试中使用最为普遍，也是客观性试题中功能最好、变式最多的一种试题。这种试题由于提供给应试者的选择答案多，通常是3～5个答案，较之其它客观性试题，具有抗干扰性强、猜测机遇率低的优点。但多项选择题又是客观性

试题中设计最难、耗费精力最多的一种试题。若要编制一套题量较大、迷惑性强、测试高层次认知能力的多项选择题，即使由一个编题技巧很高的命题班子来承担，也需花费半年甚至更长时间。

多项选择题由题干和选择项两部分构成。题干有问句、完整陈述句和不完整陈述句三种形式。选择项除一个正确答案外，其余是误道选项。现将三种不同题干形成的多项选择题分别举例于后：

例1：若 $\frac{3x+2y}{x+5y}=1$，则 $\sqrt{x+y}:\sqrt{x-y}=$ 　（题干）

（　）

A．$1:\sqrt{5}$　（误道项）

B．$3:2$　（误道项）

C．$\sqrt{5}:\sqrt{6}$　（误道项）

D．$\sqrt{5}:1$　（正确答案）

E．$\sqrt{7}:2$　（误道项）

（A—E：选择项）

例2：“九·一八”侵占东北是　（题干）

A．日本人　（正确答案）

B．法国人　（误道项）

C．俄国人　（误道项）

D．英国人　（误道项）

（A—D：选择项）

例3：下列哪一个化合物有正碘仿反应？　（题干）

A．丙酸　（误道项）

B．丙酮　（正确答案）

C．丙醛　（误道项）

D．丙醇—1　（误道项）

（A—D：选择题）

多项选择题的题型多种多样，一般随考试目的与内容的不同而变化，限于篇幅，不可一一列出，仅举六种题型为例。

（1）最佳选择题

设题者在一道试题中提出四个或五个供应试者选择的答案，要求应试者自所列选项中挑出最正确的答案。

例1：明宪宗时汪直设置西厂的主要目的是

A．侦察臣民私隐　　C．收集民间舆论

B．营造宫室苑囿　　D．巡察各地政务　　（　　）

例2：下列各因素，何者对蒸腾作用的影响最微？

A．湿度　　D．风

B．光　　E．土壤

C．湿度　　（　　）

（2）复式选择题

复式选择题是以一个图形、一个公式或一个陈述为主体，附设几种对主体的叙述，由这两部分内容组成题干。要求应试者通过对叙述与主体关系的分析和判断，再从选项中找出符合题意的正确答案。

例1：将水注入一长玻璃筒时，有声音发出。下列各项有关该声音的叙述，哪些是正确的？

①　频率逐渐增加。

②　波长逐渐增加。

③　声速逐渐增加。

A．只有①　　D．只有①和③

B．只有②　　E．①、②和③

C．只有③　　（　　）

例2：



有一列纵波形成的驻波，其质点在某一瞬间的相对位置如上图所示。

下列各项叙述，哪些是正确的？

①质点5的振幅最大。

②质点8静止。

③质点4和质点10间的距离等于波长。

A．只有①　　　　　D．只有②和③

B．只有③　　　　　E．①、②和③

C．只有①和②　　　　　　　　（　　）

（３）填充选择题

此类试题选项部分常设几个备选答案，要求应试者从中选出符合题意的正确答案，并将所选答案的字、词、字符或序号填入空白处。

例１：各门学科在知识体系上——存在某些内在联系，但各门学科的知识体系又有其独立性。

A．而且　B．所以　C．虽然　D．由于　（　　）

（４）判错选择题

判错选择题的作答要求与其它选择题相反，即要求应试者从设题者提供的几组比较项目中，选出错误的项目(不正确答案)。

例１：对健康正常的哺乳动物而言，下列各组比较项目，哪一组是不正确的？

	红血球	白血球
A．	数量多	数量少
B．	外形固定	外形不固定
C．	不具细胞核	具大细胞核
D．	较小	较大
E．	输送氧	输送二氧化碳　（　　）

（５）阅读选择题

应试者先阅读一段短文，然后根据自己的理解回答问题，即从文后的选择答案中，找出符合质问要求的正确答案。

例１：一棵树独种着，很怕风；小风吹他便要摇，大风吹他便要倒。两棵树并着种，也怕风；小风吹他还要摇，大风吹他还要倒。许多树同种着，不怕风；小风吹他不会摇，大风吹他不会倒。

（问题）：这段短文的意思是说：

A．孤独有益　　　　Б．谨慎有益

C．合群有益　　D．勇敢有益　　（　）

（6）分类选择题

分类选择题适合各种文化水平的应试者，主要用于分析、综合能力的测试。

例1．下面列举的两个词属同一类别，选项中哪一个词与这两个词的类别不同？

白、紫。

A．黑　B．黄　C．美　D．蓝　E．红　（　）

客观性试题的编制必须注重不同题型的特性和功能，恰当应用设计技巧，尤其是多项选择题的编制，更要讲究设计的艺术性．客观性试题设计的基本要领是：

（1）试题的内容和范围应严格遵循考试大纲的规定。

（2）题型选择要紧扣考试目的和考试内容。

（3）正误或是非题的答案不能模棱两可，或肯定正确，或肯定不正确。

（4）严禁抄录教材及与考试内容密切相关的书刊或参考资料上的试题。

（5）对偶题的题意与答案必须一致。

（6）同一试题选项的文字表达要长短大体一致。

（7）各道试题的正确答案应随机排列，但在同一试卷中，不同位置上正确答案的数目要基本相符。

（8）备选答案在性质上应有近似之处，迷惑性强，所设选项中，只能有一个正确答案。若选项中设有几个正确答案（如诊断式试题），要求应试者从中找出最佳答案，则必须在指导语中作明确规定。

（9）若用其它客观性试题测试应试者更适合，就不宜勉强采用多选题。

（10）试题的表述应避免暗示正确作答或提供猜测的条件。

在“电影艺术，反艺术伎俩”里，达利争论道，电影导演总的说来不能利用启开了不同于适用肉眼的自然这样的电影能力。正如沃尔特·本杰明在“机械复制时代的艺术品”(1936年)中说的：“我们的旅馆和我们的都市街道，我们的办公室和有家具的房间，我们的火车站和我们的工厂看来已把我们毫无希望地锁住了。接着又产生了电影，并且这个监狱世界被第二个十年的炸药炸开，以致于今天，在它的被远远抛弃的残垣断壁的废墟中，我们只得安静地和冒着风险地行进。”这里不是详细推测本杰明的“自由化现实”和达利的“产生于事物本身的幻想”之间的区别的地方；虽然有趣的是本杰明这样写道：“正如心理分析学说向我们介绍无意识的刺激一样，照相机向我们介绍了无意识的镜片。”

1932年，达利仍然持这样的观点，根据抽象的或发明的形式制作的电影，如象莱热或曼雷的电影，基本上是错误的，尽管他的表述更为朦胧。而现在他争辩道，电影中的诗歌仅仅在于表现“具体的非理性物”的能力，这能力只能导致名符其实的“抒情事实”。在他看来，只有在当代的喜剧影片中，才能完成这种效果。只有“具有非理性倾向的喜剧电影标志着诗歌之路”——如象马克·森勒特、哈里·兰登或马克斯兄弟(他们的《肉体的崩溃》在达利看来激起了“与雷蒙德·罗素的某些作品细节相同的抒情惊恐”。)的作品。达利对美国喜剧电影的喜爱回到了超现实主义之前的那些日子，分享这种喜爱的布纽尔在1929年与达利的一次谈话中说，波纳尔，阿道夫·门约和本·图宾比曼·雷更接近他们自己。进入《黄金时代》的热闹气氛的意外的笑剧因素在这里找到了源泉。至于商业影片，达利继续道，惊险电

例2：为什么说“知识就是力量”？

2。比较分析题。

例：请从建立制度、发展经济及提倡文化三方面，比较并说明开皇之治与贞观之治的异同。

3。陈述题。

例：客观性试题不易检测高层次认知能力的主要原因有哪些？

4。简述题。

例：简述对外开放政策的必要性及优越性。

5。求证题。

例：设抛物线$y^2=4ax$过A（1，4）、B（16，−16）两点。p点为AB的内分点，且$Ap:pB=1:4$。

①求p的坐标。

②证明p为上述抛物线的焦点。

6。例说题。

例：磷可用作弱还原剂和弱氧化剂，试各举一例说明。

7。计算题。

例：两个球形的贮藏库，容积的比为27：64。现将两贮藏库的表面涂上漆油，若较小的贮藏库需漆油72千克，则较大的贮藏库需漆油多少千克？

8。论证题。

例：为什么只有社会主义才能救中国？

9。阅读理解题。

应试者通过阅读设题者提供的一篇短文，然后根据自己的理解回答问题。

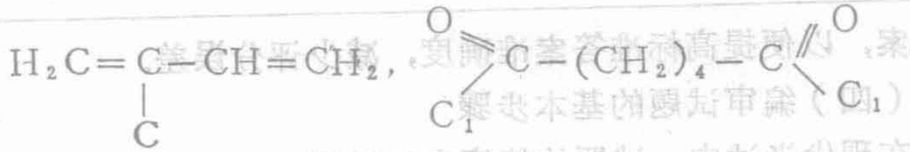
（举例从略）。

10。作文题。

例：学习、继承与创新

11。综合应用题。

例：下列三种化合物中



和

$NH_2-(CH_2)_6NH_2$

哪些可用以制造

①加成聚合物，和②综合聚合物？

试加解释，并写出这两种聚合物中重复单元的结构式．列出制造这两种聚合物的实验条件．

非客观性试题的编制尽管比多项选择题容易，但要编制一份试题数量小、内容含量大、综合程度高、如实反映欲测目标的试卷，同样是困难的。倘若欲使所编试题既不影响应试者真实水平的发挥，又能获得相对客观的测试结果，并可准确制定不同答案的分数，拟题者就需有高深的学科知识，还要熟知考试、教育、心理等方面的基本理论，有比较丰富的考试实践经验和较高编题技巧。其实，非客观性试题与客观性试题的编制各有要领，除前文已述的共同原则外，非客观试题的设计还应注意下列几点：

（1）试题欲测的应是具有科学结论或不致引起争议的知识内容，但并不排除答案的多样性。

（2）要严格控制试题为应试者提供的作答条件，以避免因线索多、条件明确而导致答案单一化，或将应试者的解题思维引向定势。

（3）试题的解答要求应简明、易懂，尽可能减少应试者为领会作答要求而花费的时间。

（4）凡需以特殊方法解答及限定作答方式的试题，应提出具体要求，甚至示以范例。

（5）同一试题包含的有关分题，以及连续质问的试题，必须逻辑严密、关联性强，不能相互矛盾或对立。

（6）每编制一道试题，就要尽量列出测试中可能出现的多

种答案，以便提高标准答案准确度，减少评分误差。

（四）编审试题的基本步骤

在现代考试中，试题的编审有四种形式、四种性质。一种是在考试机构设置诸如试题编审委员会之类的职能部门，试题编审工作由专业试题编审人员，以及选聘的部分经验丰富、熟知命题原则、具有较高试题编审能力的教师或学科专家承担，常年进行试题的编制、审定、修改、试测和筛选，属经常性质。第二种是开放性质的编制试题，即向社会公开征集试题，凡有命题能力的人都可编制、投寄试题，经专业命题机构审查选用，并付给相应报酬。或将考试大纲、双向细目表、教材、参考资料等，寄送社会有关部门事先物色的对象，由他们分散编制试题，后由专业考试部门审查、修订和筛选，从中挑选合用试题。第三种是封闭性质的编审试题，即在正式实施考试前期临时调集所需人员，组成一个一次性的试题编审班子，经过短期训练，使其明确该次考试的目的、性质、内容与对象，然后编出该次考试的试题、试卷和复本。鉴于试题保密需要，对参加命题的人员采取隔离措施，在考试实施之前不允许与外界接触，直至开考才准于“自由”。第四种是独立性质的命题，一般是由主试者个人“闭门造车”，在临考前独自编制一份试卷。此种形式多为部门内部举行的诊断性考试、研究性考试、平时测验，以及教师或研究者个人组织的考试采用。

考试试题的编制和审定，不论采用上述哪种形式进行，其试题的编审过程都要历经四个阶段：

1．明确考试要求。

明确考试要求，是试题编审的第一步。弄清考试的要求、基本性质和施测的具体对象，了解考试欲测的内容与范围，对考试大纲和试题编写蓝图规定的原则、题型类别、试题数量、测试时限、难易标准、分数比重等，做到心中有数，是这一步骤的基本

任务。

2．搜集、选择、研究编题素材。

按照不同考试目的及测试内容的实际需要，广泛搜集素材，详尽研究素材，精心选择素材，是编制各类试题的首要条件。尽管命题人员的素质和编制技巧对试题质量至关重要，但"巧妇难为无米之炊"，如若编题素材质量低劣，即使命题人员的水平、编技高超，也编不出高质量的试题。一般说来，试题素材的质量往往决定于三个方面：素材来源的宽窄；对素材钻研的程度；选择素材的方法。

试题素材搜集范围的宽窄，应以考试大纲的要求确定。比如，若要编制高校招生考试的试题，素材的收集范围必须包括下列诸方面：

（1）所考学科的教学大纲、教学计划、教材、教参资料、公开出版的试题集和复习资料；

（2）以往同类考试试卷及其统计分析资料；

（3）中学阶段教师常用的具有代表性的考试试题；

（4）各类高校系、科、专业对新生的要求；

（5）有关专家对命题内容、范围及标准的意见；

（6）国外同级同类考试的试题。

从以上方面搜集到的资料，是否能满足高校招生考试的编题需要，必须根据编题计划的要求判定，不能单以素材的类别或数量为标准，应在进一步比较、分析、研究的基础上，再决定素材的取舍或增补。

研究素材，既是熟悉素材的过程，也是对素材去粗取精的过程，大都以分析、比较两种方式进行。所谓比较，系指纵横交叉比较。纵要比较不同年份、不同时期同类考试试题的优劣，不同阶段教学计划、教学大纲、教材内容的变化，以及各类高校、系、科、专业对新生要求的变化等。横要比较不同地区教育水平的差

异，不同国家同级同类考试试题难易度的高低，各专家意见合理成份的大小等。素材分析，包括教学计划、教学大纲和考试大纲的分析，进一步明确素材选择的要求；教材内容的分析，以了解重点与结构；历年试题利弊因素分析，以吸取成败之经验教训；同类考试试题分析，以明辨长短、合理吸收；专家意见分析，以分辨异同，权衡利弊、择善采纳。经过这样的分析和比较，即可转入素材选择。

选择试题素材需态度慎重、方法科学。这一环节应注意四点：

（1）所选材料要有针对性，既要符合被测对象的实际，又要反映考试欲测的内容；

（2）选取的材料应重点突出，能体现所考学科教学的主要目标；

（3）素材的种类齐全、数量足够，凡考试大纲规定测试的内容，不管在考试总体内容中比重大小，都要有充足的素材，不致编写试题时缺材。

（4）教材应具有代表性和公平性，准备用于编制试题的素材，当是被试者必须掌握的知识和理应具备的技能及能力方面的材料，用所选素材编制的试题，必须使参加考试的每一个人都可获得大致相同的应试条件。

3．编制试题。

选齐了编题所需要的素材后，就可进行第三步的工作，即运用选定的材料、按编题计划要求着手试题的编写。

现行考试的试题编写，一般是采取分散与集中相结合：一方面，通过分头编写使命题人员各扬其长；另一方面，通过集体讨论互补其短。具体作法是，根据各编写人员的特长，将编题任务作恰当分工，或一人承担一部分，或各编一套。任务确定后，先由编写人员遵照统一制定的编题计划、原则和共同选定的材料分

头编写，然后将各自编写的试题集中进行讨论，共同选出送审试题。出于集中讨论和审题的需要，各命题人员编写的试题，必须按规定抄写在统一制作的命题卡上，内容包括所属学科、试题原文、正确答案、测试目标、试题编号、命题人姓名、编题时间、审题意见、审题人姓名、审定时间、分数比重等，图7－8、图7－9是一道多选题的命题卡式样。

4．试题审查与初选。

对试题进行审查和初选，是编制原始试题（未经预测检验的试题，又称初题）的最后一道程序。这一步骤的任务是，对命题人员编出的原始试题进行一次初步筛选，将符合考试大纲要求、达到设计规定标准的试题挑选出来，重新分类编号，作为预测的备用试题；对修改后才能入选的试题提出详细意见，退给原命题

所属学科		题　　号	NO：
测试目标			
试题原文			
正确答案		分数比重	
编题人		编题日期	年　月　日

图7－8　多项选择题命题卡式样（正面）

审题意见			
审 题 人		审题时期	年　月　日

图 7－9　多项选择题命题卡式样（反面）

人进行修改，并以同样程序再行审查；对不符标准的试题予以废弃，但必须理由充足，不可轻易否定一道试题。

审题由两种人组成的审题班子承担。一种是考试机构内部的专职审题人员，他们大都是学科考试的主要负责人，长期从事试题评审工作，一般不更换职位。另一种是兼职审题人员，由考试机构外部的学科专家或在职教师充任。兼职审题人员不长期固定，常常是根据试题审定的实际需要雇请。参加试题审查的专、兼职成员，均有严格的要求，必须具备下列条件：

（1）应是被审试题所属学科的专家，考试理论的造诣较深；

（2）能熟练应用审题的原则和技巧；

（3）对试题质量的优劣有较强的鉴别能力；

（4）具有高度的责任感和优良的考德。

五、预　测

预测作为检测、筛选试题的一种手段，是现代考试提高试题质量、保证考试结果有效可靠的重要措施，也是现代考试设计中的关键步骤。但它并不是人类考试设计历来就有的程序，在传统考试的设计中，试题的筛选是在试题设计的第四步骤中（审查、初选试题），以主观判断的形式完成的，至于一份试卷难易度和解题的所需时间检验，也并未抽样测试，多数是由拟题者亲自做一遍（有时请人做一遍），然后根据试做情况，决定试卷的份量和正式考试的时限。此种测验手段，与现代考试设计中的预测有着根本区别。同时，即使在现代考试中，预测也不是所有考试设计必不可少的步骤，比如部分系统举行的安置考试、研究性考试、文化普查考试等，就无需考前预测。现行一些大规模的、具有一定社会影响的和较为正规的成就、合格、选拔性考试，之所以把预测作为考试设计的重要步骤，关键在于预测是通过考试实践鉴别试题质量，对试题的检测和筛选比主观判断客观准确。

（一）预测的任务和目的

预测的任务，是对初步审查合格的原试试题进行实测检验，看初选试题的内容是否反映了欲测的目标？难易程度是否适合正式考试对象的实际水平？试题的鉴别能力是否达到规定标准？指导语是否能为应试者准确理解？试题类型、各类试题的比重、全卷份量、考试时限等方面的规定是否恰当？测试的实际效果能否达到预定标准？预测有三方面的目的：对经过审定且认为合格的试题，在客观检测的基础上再作筛选，为正式考试挑选真正符合标准的试题，以保证正式考试的可靠性和有效性；获取施测规范化的经验，以便正式施测时有效控制与考试目的无关因素的干扰，减少测试误差；全面分析测试结果，为建立同类考试常模获

取依据。

（二）预测的方法

预测的方式和方法，当视具体条件和环境而定，以能够客观检测试题质量、获得真实准确的反馈信息为原则，不存在固有的模式。现阶段使用比较普遍的预测方法有三种：

1．设置预测区。

这种方法是在某一区域或一个国家范围内，选择一个具有代表性的地区作为预测区。预测区，既可为一次性考试的预测而设置，也可相对固定，作为某一时期一种或几种考试的预测区。这种预测方法组织、实施便利，其缺陷在于试题难以保密，容易影响正式考试的真实性，不宜多用。

2．分散埋伏。

这是目前采用最多的预测方法。分散埋伏预测有三项具体措施。第一，把试题拿到与将来正式考试对象具有大致相同条件的人中间进行预测，亦即通过随机或分层抽样的方法组织预测。第二，提前一年甚至几年时间，把初选试题分散安插到与将来正式考试对象同类人员的平时作业中进行预测。采用此种方法预测，不易控制测试环境、时间，很难获得完整准确的反馈信息。第三，把试题分散放在同类考试的正式试卷中，作为不计分试题进行预测。这种预测方法，对确定正式考试的时限、试卷长度等，不能获得可靠依据。

3．聘样本团。

这是一种规模小、涉及范围窄、保密性较强的预测方法。基本作法是：在试题编审完毕，并根据编题计划拼出预测试卷后，聘请一批与正式考试对象条件相等而又不参加正式考试的人，作为预测的团体对象，然后按正式考试的程序和要求，对他们进行一次或几次测试，以便对预测试卷或等值试卷作全面检测，为正式考试的实施取得多方面的依据和经验。但是，聘样本团的预测

方法有其严重缺陷：样本数量受限，不利于项目分析；如果预测试题无题库存放，势必在近期举行正式考试，否则样本团的成员就没有与外界接触的“自由”。

上面例举的三种预测方法，实际效果各有利弊。相比之下，在较大范围内通过随机或分层抽样预测的方法，适合于增补题库试题的检测；将试题安插在正式考试试卷中（计分与不计分均可）预测的方法，对检测某些周期循环举行的考试（不用题库试题）试题比较适宜；通过聘样本团进行预测的方法，有益于某种关系重大的一次性考试试题的检测；设预测区的办法，对设计某种新考试的试题检测更适合。因此，选择预测方法，应根据实际条件和客观需要决定。

（三）预测的基本要求

预测的方法可灵活多样，要求却是相同的。预测既要收到客观检测试题质量的效果，又不能给正式考试造成不良影响。具体要求是：

1．对象要有代表性。

预测，实际上是正式考试的模拟测试，参加预测的应试者（预测对象或测试样本），在年龄、知识量、认知能力、健康状况等方面，应与将来正式考试的对象大体一致。或者说，预测对象的总体结构能基本代表正式考试对象的总体结构。否则，就不成其为样本或样本团，预测也就没有实际价值。按理说，要使预测对象确有代表性，就得从符合预测条件的总体中抽取应试对象，总体分布的范围有多大，取样的范围就应有多大，但客观上难以办到。譬如我国的普通高校入学考试，若要在全国范围不同水平的高中应届毕业生中抽样预测，是不太现实的，只能从一些比较有代表性的省市中抽样挑选预测对象。当然样本数量应尽可能大一些，否则，即使确有代表性的对象参加同一份试卷的预测考试，也不能为试题质量的分析提供足够的有效数据。另外，对于

预测试题的效果来说，预测对象取样的范围愈大愈可靠，但必须以切实有效的保密措施为前提，不然会出现取样越广泛，对正式考试副作用越大的后遗症。

2．应试条件要与正式考试近似。

所谓应试条件，是指测试的内容、考试的组织形式、施测程序、应考规则、应考环境和考试时限等。在这些方面，预测的考试时限可适当放宽，尽量让应试者做完全部预测试题，以便获得全面系统的信息资料，进行项目分析时有丰富的素材。除此之外，其它各点均应与正式考试的条件基本相同。

3．必须强化被试的主体意识。

考试是主试与被试两种主体的交互活动，任何考试结果，都是主试与被试双方交互作用的产物。若考试过程中两种主体配合不密切，将直接影响考试结果的真实性。由于预测对应试者个人来说，与其切身利益并无利害关系，压力不大，动力缺乏，是应试者普遍存在的现象。要使他们在预测中积极主动地给予配合，考前主体意识的强化十分重要。为此，考前应做好三件事：说明预测意义，使应试者明确参与预测的责任，形成良好的应考心理；给予一定时间，指明复习范围，敦促应试者认真备考；进行必要的应考训练，使他们在预测时不致逾距，保证预测按既定规程进行。

4．测试内容应高度保密。

在预测的各种措施中，保密措施处于头等重要地位。它不仅关系到预测目的可否实现，而且涉及到正式考试是否有效的问题。如前所述，预测的根本目的，是为正式考试挑选高质量的试题，获取决定正式考试试卷长度、难易指标、测试时限及其它各个方面的依据。如若缺少严格的保密措施，测试内容被泄露，即使预测选出的试题是高质量的，但作为正式考试用题却是无效的。因为试题一旦泄密，它在反映学生水平的真实性方面就失去

了固有效用，以这种试题作为正式测试的内容，测得的结果必然是虚假成份大于真实成份，不具客观标准。所以，若用已经泄密的预测试题施考，既不科学也不公平。但是，对于一些需在较大范围广泛抽样进行的预测，要保证不泄露测试内容，的确不是一件容易事。从现今国内外所采用的各种保密措施来看，还是题库保密法较为科学，即提前几年进行预测，然后把合标准的试题存入题库，几年后再从题库中提取正式考试的题目。采用题库保密法，需具备两个前提：首先，从预测到正式考试这一时期内，同级教育的水准应相对稳定，没有大起大落的现象，预测对象与正式考试对象的实际水平无显著差异。其次，预测应提前较长时间，至少要比正式考试提前1～2年。总之，当年预测的试题不能用于当年的正式考试，这是基本原则。

5．测试动态要有详细真实的记录。

为正式考试施测过程的质量控制获取经验，是预测的重要目的之一。倘若预测是在拥有录像设施的专用试场进行，这种控制经验的信息可从录像中获得，但对于绝大多数的预测考试来说，施测过程的质量控制信息，只能从预测考场动态的真实记录中获得。因此要求预测应在与正式考试同等条件和规格的环境中进行。另一方面，又要求主试除在施测过程中防止考生作弊外，还必须细心观察考生的应试情况及测试的进展情况，并对考生应试时的各种情绪反映，解题速度，在测试过程的不同时间段考生答完了哪些试题，考生对指导语、题意表述的理解情况，整个测试实际占用多少时间等，作真实详细的记录，借以实现预测两个方面的目的：将现场记录与预测结果的统计分析资料相结合，对预测试题进行全面的综合分析，提供筛选试题的质量；通过对现场记录的分析，从中找出施测环节影响考试信度、效度的因素，为正式考试的实施误控经验，便于有的放矢的制定误控措施，在正式考试的施测过程中，更好地防止与测试目的无关因素的干扰。

（四）预测的三大步骤

预备测试的全部流程，近似于正式考试的实施流程，可细分为制订预测实施方案、拼配预测试卷、施测、结果分析、修改和调整试题等多个环节。若按预测各阶段的主体任务划分，也可分为三大步骤。

1．试前准备与制卷。

预测的试前准备工作与正式考试有所不同，它是由命题人员和考务人员共同承担，而不是考务一个部门负责。这一步骤的主旨，是为正式预测作准备，其任务包括制订预测方案、挑选预测对象（确定预测区或聘请样本团）、设置预测试场、拼配并印制预测试卷及复本。如果条件允许，在正式预测前，还可把准备用作预测的试卷先试测一次，即挑选少量与预测对象条件同等的人为样本，在他们中间进行一次初步测试，并对测试结果作简单的统计分析，若有问题，还可进一步修改、调查，以增强预测试卷的实际效用。

2．正式预测。

一旦准备工作就绪，可随即进行正式预测。有关正式预测的方式、方法、任务和要求，前文已分别介绍，不再详述。尚需强调的是，凡是为一次性考试或创设某种新的考试而组织的预测，为确保正式考试的质量，或新设计的考试实施圆满成功，不能仅用一份试卷对同一样本团体一次预测的结果，作为决定将来正式考试试卷标准的依据。要么用同样的试卷对几个样本团体预测，要么用相同的试卷对同一样本团体先后测试两次，使预测提供的数据和信息更为准确。

3．统计分析预测结果。

这一步骤是预测的中心环节，其主要任务是，借助电脑技术，采取定性与定量相结合的方法，对经过预测的试题作全面系统的分析，了解每道试题的难易度和鉴别力，检查试卷的适应

性、可靠性和有效性，如果没有明显问题，再根据统计分析的结果，筛选出正式考试的试题，确定或调整正式考试试卷的难度；区分度指标、内容份量、测试时限、题型及比例。最后将筛选出来的试题重新编号填卡，或直接用作正式考试的制卷，或存入题库作为今后同类考试的备用题。

试题和试卷质量的统计分析，是一项复杂细致的工作，有关统计分析的内容、任务与方法，将在本章第二节作专门讨论，此处从略。

六、编制标准卷与等值复本

编制正式考试使用的标准试卷，制作与标准试卷同等效用的试卷复本，二者在要求上是一致的，即符合考试目的、能准确测出应试者被测方面的实际水平。标准试卷与等值复本的制作若是采用题库中的试题，则应放在考试实施的第二步骤中进行。若考试设计无预测环节，在试题审定工作结束后，即可着手此项工作。如果是将预测试题直接用于正式考试（不入试题库），作完预测结果的统计分析后，就可转入标准试卷和等值复本的编制阶段。该步骤的主要任务有三项：

（一）编制标准卷

编制标准试卷分两步进行。第一步，从经过预测和统计分析合格的试题中挑选需用试题，并把选出的试题重新审查一遍，看在内容、数量、题型类别、不同类型试题的比重、难度等方面，是否与考试大纲的要求及双向细目表规定的指标相符，若有出入，应再行调整，确无不符之处后，方可转入编制工作。第二步，根据双向细目表规定的各项指标要求，将挑选出来的试题进行合理拼配，编出供正式考试用的标准试卷。

编制标准卷的核心问题，是试题的编排与组合，亦即如何使试卷的结构合理。一般说，从经过预测的试题中再选出的试题，

其质量是可以置信的，但在拼配标准卷时，对可靠有效的试题编排组合不当，正式考试同样达不到预期的信度和效度。从长期制卷的实践经验看，标准试卷的制作必须注重四点：

第一，在考试内容、题型、各类试题的比例方面，必须紧扣编题计划；

第二，在试题的难易程度及编排序次方面，必须适合测试对象的实际水平和考试应试的心理特征；

第三，试卷长度方面，必须考虑测试的时限；

第四，试卷拼配就绪后，还必须对照考试大纲和双向细目表再核定一次，所制试卷确与要求相符，才能定为正式考试的标准卷。

（二）编制等值复本

等值复本是一种与标准试卷等值的试卷，人们也称之为副卷或复份。当标准试卷制成后，再编制几份甚至数份等值的试卷复本，有着多方面的作用：当标准试卷出现泄密或发生其它问题时，可及时起用等值复本，保证考试的正常进行，弥补因试卷问题造成的损失；某些规模宏大的考试，如果采用多份等组试卷同时施测，可获得更为全面丰富的信息，能提高考试的质量和增强考试的实际效用；同一试场同时用几份等值试卷施测，能较好地防止考试舞弊；对等值复本进行再次预测，其测试结果，可作为建立同类考试常模的依据。

编制等值复本的要求，与标准试卷无异。它的制作，必须与标准试卷依据同一份考试大纲和编题计划。所谓等值，系指复本应与标准试卷具有同等效用：适合同一考试目的；有同等的难度、区分能力、信度和效度；试卷长度、考试时限、测试对象、测试内容、实施程序、评分及分数解释的方法等，都是相同的。只有一点例外，即复本所用试题，与标准试卷、其它复本或同类试卷不能相互重复。当然，各份试卷的等值是相对的，只能做到

标准基本一致，不能说绝无差异。要编出两份或多份与标准试卷绝对等值的复本，事实上做不到。

编制等值复本的程序及方法是：先按拼配标准试卷的同样要求，从经过预测的合格试题中选出所需数量的试题，再把试题按难度大小以1、2、3、4、5、6、7、……的顺序依次排列，直至试题全部排完为止，然后根据欲制等值复本的数量，用以下方法进行分配与组合：

1．两份等值复本编制法：

A卷：1，4，5，8，9，12，……

B卷：2，3，6，7，10，11，……

2．三份等值复本编制法：

A卷：1，6，7，12，13，18，……

B卷：2，5，8，11，14，17，……

C卷：3，4，9，10，15，16，……

若要编制更多的等值复本，方法同上，以此类推。为检验编制的复本是否与标准试卷等值，在条件方便的情况下，还可把等值复本再作一次预测。这样更为可靠。

（三）制定标准答案

标准答案是鉴别应试者答题优差、正误的依据，判定分数的统一标准，也是提高考试信度的重要措施。在结束标准答案和等值复本的编制工作后，应精心设计一份标准答案，确定各道试题评分的标准及分数比重。如果设计的是客观性考试，还要制作统一的答卷纸。

作为鉴别考生答题正误唯一准则的标准答案，必须符合标准客观、准确无误、表意确切、对错分明、易于掌握、使用方便的基本原则，评分时不受主观因素的影响，不管是手工方式评卷还是电子计算机评卷，都能对应试者的测试结果作出真实客观的评价。

客观性考试的设计，之所以要在拼配标准卷的同时作统一的答卷纸，是由于客观性考试的答卷形式与传统考试有别。为便于机器阅卷，客观性考试通常把试卷与答卷分开。在答卷纸上，按照随机或序次排列的方法，列出各道试题的序号和选项符号（如图 7-10）。施测时，应试者只需按统一要求涂黑各题选次符号中自认为正确的符号。试卷的评阅，由光学扫描器阅读考生答卷，将考生答题的信息准确输入电子计算机，以判明正误，并算出每个考生应得的分数。

图 7-10中例举的“1，2，3，4……”代表试卷上的试题序号；Ⓐ Ⓑ Ⓒ Ⓓ 为各题选择答案的符号；考号后面的6个方框，供考生填写本人的准考号码；考号下面的6排数码供考生为自己的考号作标号用，以便光学扫描器阅读。设计答卷纸时，考号后面的方框应与考号下面的数码排数相等，方框的个数和数码的排数，根据考试实际使用的最后一个考号数字确定。考生在正式答卷之前，除将考号填入方框外，还要按考号的序次，分别涂黑各排数码中与考号相同的数字。假设某考生的考号为023578，就先把这个号码填入考号的方框，然后在第一排数码中将⓪这个圆圈涂黑，接着在第二排数码中涂黑②这个圆圈，依次类推，一直把与考号相对应的6个数字全部涂黑为止。

七、建立常模

建立常模的任务，是为同类或某次考试制作标准量表，以此作为平衡试卷的标准、转换考试分数的参照点和解释分数的依据，亦即制定同类考试或某次考试的标准，为计分和分数解释客观化创造条件。没有常模，同类考试分数就没有可供比较的统一标准，分数的解释无据可依，同一分数的含义也就不能得出公认的结论。因此，人们把常模称为客观性考试的生命。

（一）常模的含义

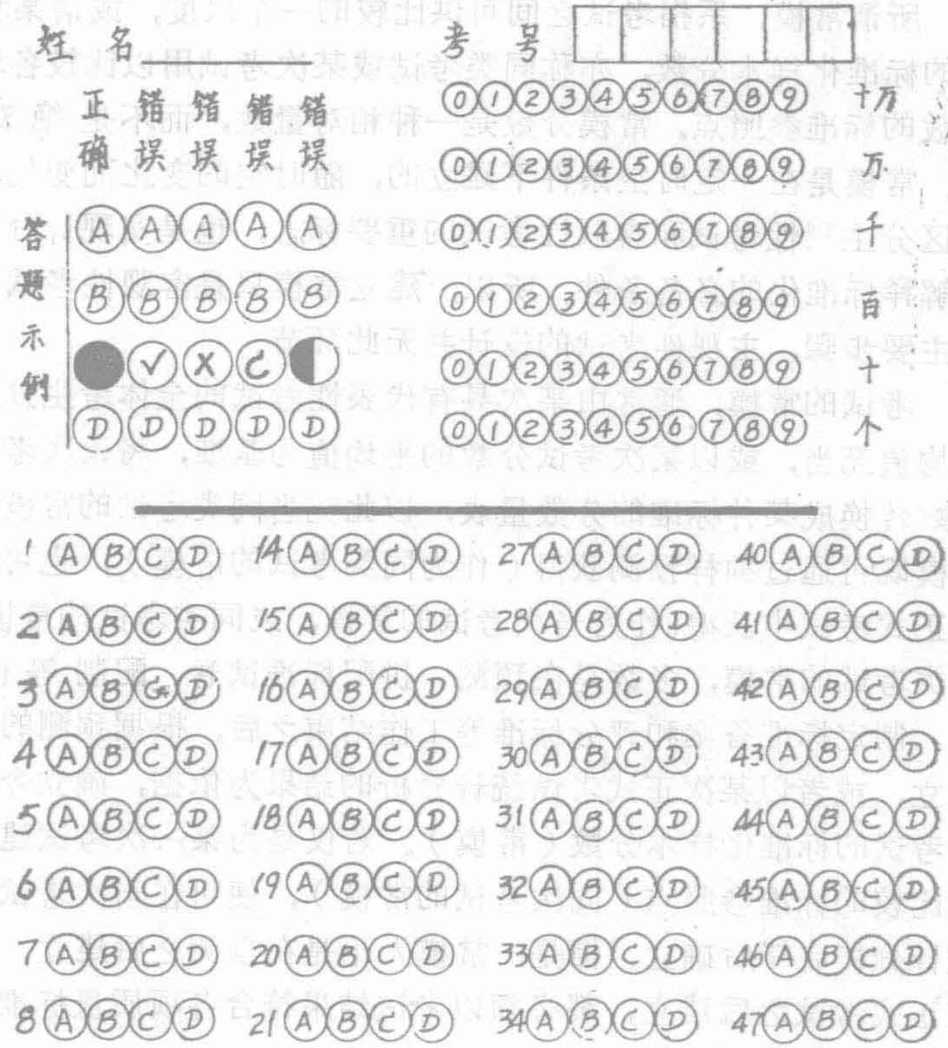


图7-10　客观性考试选择题答卷纸式样

所谓常模，系指考试之间可供比较的一个尺度，或谓某种考试的标准化样本分数，亦称同类考试或某次考试用以比较各考生分数的标准参照点。常模分数是一种相对量数，而不是绝对量数。常模是在一定时空条件下建立的，随时空的变化而变化，它是区分主观性考试和客观性考试的重要标志，也是实现计分和分数解释标准化的必备条件。所以，建立常模只是客观性考试设计的主要步骤，主观性考试的设计并无此环节。

考试的常模，通常由某次具有代表性考试的全体考生分数的平均值充当，或以某次考试分数的平均值为基准，将该次考试的分数转换成某种标准的分数量表，以此充当同类考试的常模表。常模既可通过抽样预测获得（作为同类考试的常模），也可直接从正式考试中获得(作为当次考试的常模，或同类考试的常模)。同类考试的常模，多数是在预测、拼配标准试卷、配制等值复本、制定标准答案和评分标准等工作结束之后，根据预测的结果建立，或者以某次正式考试统计分析的结果为依据，确立今后同类考试的标准化样本分数（常模）。若仅是为某一次考试建立分数比较的标准参照点（同次考试的常模），便可在当次考试结果统计分析完毕后确立。但是，常模无论是在预测之后建立，还是在正式考试之后建立，都必须以考试结果符合各项质量标准、确系可靠有效为前提。

（二）常模种类、功用及求常模之方法

在考试或测验中，人们为比较各次考试质量的优劣、衡量某级教育水准的升降、准确解释考试分数的意义，通过不同方式，建立了多种类型的常模，诸如年级常模、年龄常模、标准分数、T分数、教育商数、成就商数、百分等级、标准阶段等。现就考试中使用较普遍的几种常模介绍如下：

1．年级常模。

年级常模又称级模，含义为各年级的平均成绩。在年龄常模

时兴之前，它是学校教育考试最常用的一种常模。建立年级常模的基本方法是，对不同地区同级学生进行一次某种统一的考试，再统计分析考试结果，求出参加该次考试全体考生成绩的平均值（平均分数）。这个平均分数就是该年级的常模。其计算公式为：

$$\overline{X}=\frac{\Sigma X}{N} \tag{7·1}$$

(7·1)式中：$\overline{X}$为某次考试全体考生所得分数的平均值。

ΣX为该次考试全体考生所得分数之和。

N为参加考试的考生数。

如果在某次五年级学生的语文考试中，全体考生的成绩按公式(7·1)计算，其平均分数为50分，那么，就可以把50分定为五年级的语文考试常模。

为使用的便利，也可改变一个年级仅建立一个常模的作法，即把一学年分为10个学月，在同一学年定出几个可供比较的常模，以检验同级学生在不同学月的学业成就。美国麦柯尔制作的G分数，属于此类年级常模。某年级学生在某学科考试中的平均分数，若与某年级某学月考试的常模分数相等，便表明这个年级的学业成就已达到某年级学习几个月后的平均水平。假设在一次数学考试中，五年级学生的平均分数为25分，这个分数恰与五年级第3个月学习成绩的常模分数相等，说明该年级的数学成绩已达到五年级学习了3个月数学的平均成绩。值得注意的是，这种方法“是假定学生在一学年内，每一学月的进步是很有规律的”。当这种假设与实际情况不符时，对考试分数的解释就不可能准确。

年级常模具有求取简单、易于理解、使用方便的优点，适合于小学教育各学科的成绩考试，对高层次教育的学业成绩考试效用不大。

年级常模的主要缺点在于：

（1）年级常模不能准确判定每个学生学业成绩在其总体中的相对地位。因为年级常模是从考生原始分数中获得的一个平均分数，各次考试的原始分数没有同等的值，各次考试的分数不可比。

（2）由于年级常模只代表各年级学业成绩的平均水平，而并不反映每个学生都能达到、都达不到或只能达到这个水平。所以，容易产生重视学生平均成绩、看轻个别差异的弊端，进而影响个性的培养。

（3）年级常模是以一学年的平均成绩为参照分数，去衡量同年级学生的学业成绩。所以使用年级常模的学校必须有相同的升留级制度、教学进度和课程设置。如果一批学校是以学生年龄为标准决定升留级，而另一批学校学生的升留级是按学习成绩的优劣决定，或者各学校课程设置的年级各不一样。在这种情况下，检验各学校同年级的学业成绩，就不能以年级常模为尺度。

（4）年级常模不能与智力测验结果（智力年龄）作直接比较。

2．年龄常模。

年龄常模，即是各年龄儿童的平均成绩，或谓各年龄儿童教育成绩的平均分数，所以，又称年龄常模为教育年龄，它表示各年龄儿童的教育程度。求年龄常模的方法，一般是对各年级儿童进行一种或几种教育测验，再用统计方法求出各年级测验分数的平均值，以此作为各年龄儿童的常模。下面以斯坦福学业测验年龄常模表（表7－5）为例，进一步说明年龄常模的求法。

表7－5　　学业测验年龄常模表

年　龄	9	10	11	12	13	14
平均分数	23	33	46	57	66	72

表7－5中与年龄相对应的平均分数，就是该年龄儿童学业成绩的常模。比如66分，就是13岁半儿童的常模（因为13岁包括满了13岁而未满14岁的儿童，其平均年龄为13岁半）其他年龄的常模余类推。至于两个年龄之间的儿童，如11岁8个月，11岁11个月的儿童，其年龄常模则采用推算法求出。

年龄常模是以年龄为单位计分，相对年级常模，它有分数稳定不受学级编制的影响，可与智力年龄直接作比，也可与其它教育测验相互作比的优点，并能反映同龄学生学业进步的速度，以及不同年龄儿童的学业成绩在各年递增的量度。但年龄常模的缺陷也客观存在。较为突出的是：

（1）年龄常模是由抽样测试获得，其常模标准难易受测试样本实际水平的左右，很难求得可靠公允的常模标准。

（2）年龄常模是以平均分数衡量儿童的学业、达到某个教育年龄。但一个儿童的教育年龄，是多因素的集合，即使两个同龄儿童的教育年龄相等，也并不等于两个儿童构成教育年龄的因素组合相同。单以平均分数衡量一个儿童学业成就的大小，每个儿童的特殊才能及对教育的特殊要求容易被忽略，不利于因材施教，影响个性的发展。

（3）由于儿童各年龄间学业成绩的增长不是等量递增，所以年龄常模的计分单位相同而不相等。

（4）在使用中，年龄常模以年龄为单位反映学生的学业成绩，人们不易适应。

（5）适用范围较窄，只适宜于小学教育阶段使用。

3．标准分数常模。

标准分数常模，亦称标准分数量表或Z分数量表。它以平均数为参照点、标准差为单位，反映一个学生的学业成绩在同一考试总体中所占的地位。所以，标准分数常模是一种相对地位量数常模。

在统计学中，标准分数称为Z分数，一般将其定义为，某一数列中的某一原始分数与其平均分数的差数除以标准差所得的商数，以公式表示即：

$$Z=\frac{X-\overline{X}}{S} \tag{7·2}$$

(7·2)式中：Z为标准分数。

X为某原始分数。

$\overline{X}$为某一列原始分数的平均分数（$\overline{X}=\frac{\sum X}{N}$）。

S为标准差（$S=\sqrt{\frac{\sum(X-\overline{X})^2}{N}}$）。

标准分数的平均值为0，标准差以1为单位，标准差的范围在－3～＋3之间，其分布形式如图7-11所示。

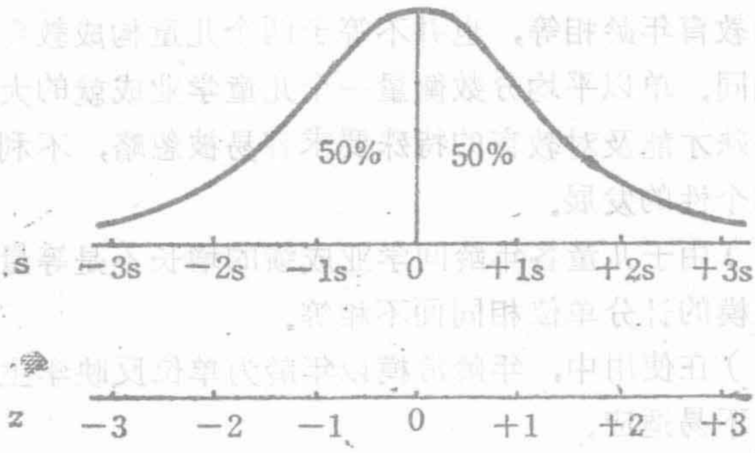


图7-11　标准分数常模示意图

自图7-11中可看出，标准分数的正负数是表示大于或小于平均数，如果用以表示考试成绩，其中正数表示高于平均成绩，负数表示低于平均成绩，0等于全体考生的平均成绩，即比他成绩优和差的考生各占50%。例如，我们已知某考生化学考试的原

始分数为84分，化学考试全体考生的平均分数为48分，标准差为12分，用公式(7·2)求该生的标准分数：

$$Z=\frac{X-\overline{X}}{S}=\frac{84-48}{12}=\frac{36}{12}=+3$$

计算结果表明，该生的标准分数为＋3分，高出全体考生平均成绩3个标准差，或者说，该生的成绩优于99.9%的考生。

标准分数是使用最普遍的一种常模参考分数。由于它是一种相对的地位量数，数值抽象，单位相等，可将不等值的分数转变为等值分数，使单位不等的考试分数能比较、加减或平均，也可与百分等级、智商等常模分数直接比较，如图7-12所示：

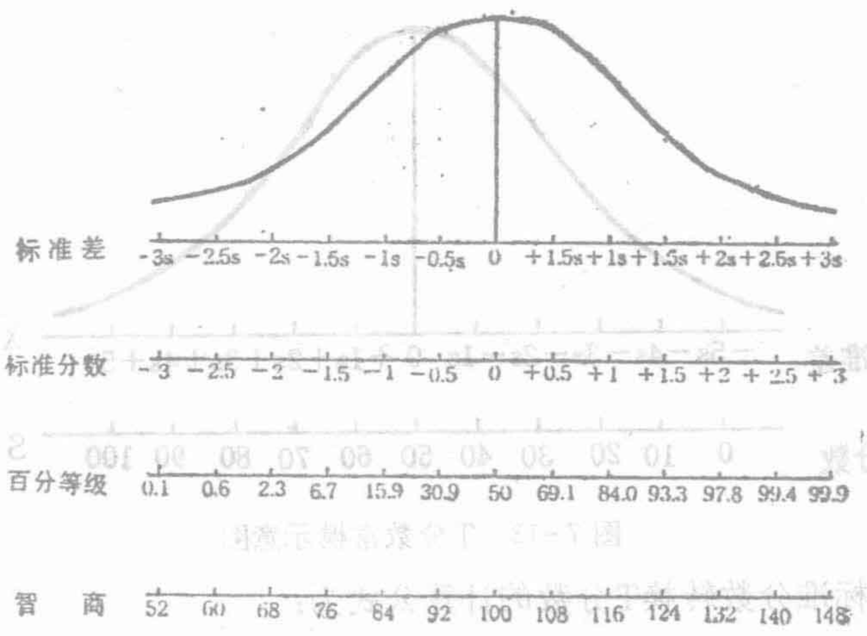


图7-12、三种常模分数比

使用标准分数常模的缺点是，计算甚为不变，考生成绩常小数出现，且有正负之分，不易理解。

4．T分数常模。

T分数常模是标准分数常模的变换形式，又叫T分数量表。

它同样以平均数为参照点，以标准差为单位，反映一个学生的学业成绩在同一考试总体中所占的地位。但T分数是经过线性变换的一种标准分数，它与前文介绍的标准分数有所不同：标准分数是对服从标准正态分布的原始分数的转换，T分数是对服从标准正态分布的标准分数的转换；在标准化常态分配中，标准分数的平均分数为0，标准差为1，平均数以下3个标准差为起点，而T分数的平均数为50，标准差为10，它以标准差的1/10为单位，平均数以下5个标准差为起点，整个分数分布在－5～＋5的范围　　于平均分数5个标准差或大于平均数5个标准差的考生分数，仅　0.00003，参见图7－13。

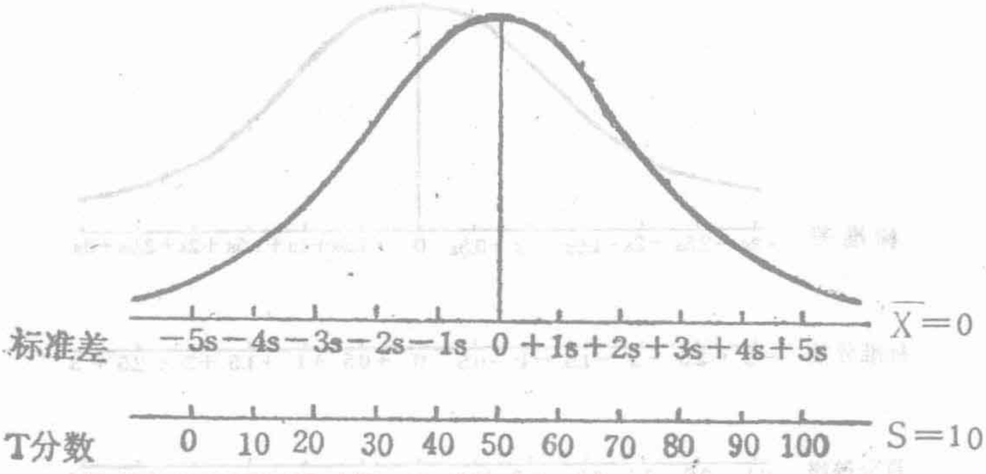


图7－13　T分数常模示意图

标准分数转换T分数的计算公式为：

$$T = 10Z + 50 \qquad (7\cdot3)$$

(7·3)式中：T为T分数。

　　　　Z为标准分数。

求出考生的Z分数后，转换成T分数十分便利。假设某考生的标准分数为＋2.5分，若要转换成T分数，其计算程序及结果如下：

$$T=10Z+50=10\times2.5+50=75$$

若要直接将原始分数转换成T分数，计算公式则为：

$$T=10\frac{(X-\overline{X})}{S}+50 \tag{7·4}$$

(7·4)式中：T为T分数。

X为某个原始分数。

$\overline{X}$为某一列原始分数的平均值。

T分数的标准差和平均数，根据实际需要确定，即任选两个常数充当。因此，通过变换平均数和标准差的办法，可建立多种形式的T分数常模。如英国学校考试使用的T-分数的平均数为50，标准差为10；美国大学入学考试委员会使用CEED分数的平均数为500，标准差为100；美国陆军普通分类测验使用的AGCT分数的平均数为100，标准差为20；我国EPT考试常模分数的平均分数暂定为90，标准差为20；MET考试的常模分数，暂定平均分数为60，标准差为12。其变换形式如图7-14所示。

使用T分数常模的优点在于：首先，T分数把最小分数定为0，最大分数定为100，其分数的形式与普通百分制无异，使用时人们易于适应。其次，T分数以标准差的1/10为单位，不仅单位相等、准确可靠，而且可反映每个考生的成绩在其总体中所占的位置，即每个考生的成绩在考试总体中超过或低于其他考生成绩的人数比例，比如，某考生的T分数是70分，只要通过查T分数与百分比对照表（见附录表1），就可知同一考试总体中有97.72%的考生成绩比他差，有2.28%的考生成绩比他优。第三，T分数常模可与其它标准分数常模直接比较，适用范围广。

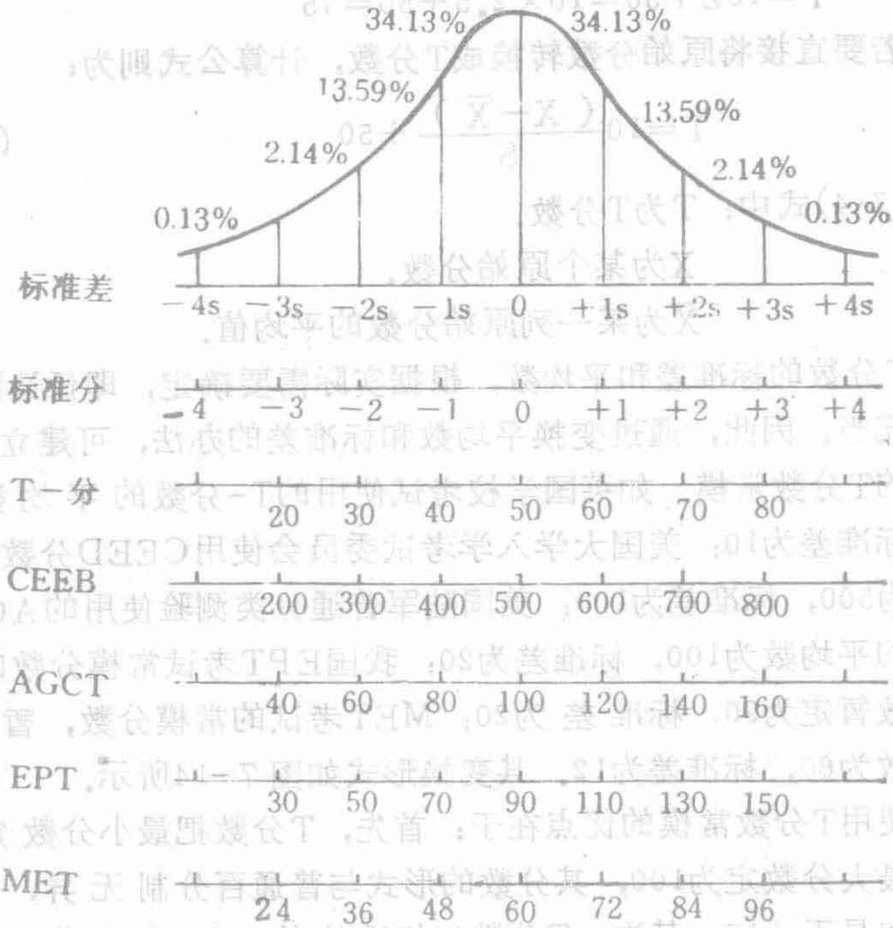


图7-14 几种平均数和标准差不同的T分数常模示意图

5．百分等级常模。

百分等级常模，是用百分等级法制定的一种百分位置表。它以100代表考生整体，不论参加考试的实际人数超过还是少于100，均按100个等级划分，即把一次考试的全部原始分数转换到100个等级范围的常模表（百分位量表）上，获得最高考试分数的考生居于常模表100的位置，获得最低考试分数的考生居于常模表0的位置，常模分数的平均值居于50%的位置。

百分等级是一种相对地位量数，它表示某一量数在所属团体中所超过的单位数占总单位数的百分数。如果某考生在生物考试中的原始分数转换后，在百分位量表上居于87的位置，便表明该

考生生物考试成绩的百分等级为87，说明全体考生中有13%的人比他的成绩高，有87%的人比他的成绩低，或者解释为该生的生物考试成绩要优于87%的考生。同理类推，某考生考试成绩百分等级的量数越大，表明他的成绩在考试总体中的位置越领先，百分等级的量数越小，在考试总体中的位置就越偏后。

但必须明确，百分等级的量数与我们平常使用的百分制原始分数不同，按百分制评出的原始分数，它只表明考生得分数量的多少，不反映每个分数在总体中所占的位置。此外，百分等级的量数，也不同于人们使用的一般等次概念，人们平日按得分多寡依次排出的第几名，它只表明分数大小的先后次序，也不能反映某个名次考生成绩在总体中的位置。百分等级量数它所表示的是，当每个考生的原始分数转换为百分等级后，它在同一考试总体中所占的地位（位置）。下面介绍两种求百分等级的方法。先以某市376人英语考试的百分位表（表7－6）为例，说明分组原始分

表7－6　　某市376人英语考试百分位表

组　距	次　数	累积次数	百分位
60—64	2	376	100
55—59	12	374	99.5
50—54	20	362	96.3
45—49	32	342	90.7
40—44	46	310	82.4
35—39	58	264	70.2
30—34	64	206	54.8
25—29	58	142	37.7
20—24	42	84	22.3
15—19	23	42	11.2
10—14	15	19	5.0
5—9	4	4	1.1

数百分等级的计算方法。

表7－6各栏数值的含义是：第1栏（从左至右）的数值表示各组原始分数的组限，第2栏的数值，表示在相对应的原始分数组限内考生原始分数分布的次数，如2就是60-64这一组限内考生原始分数的分布次数；第3栏的数值，表示所在组以上各组考生原始分数分布次数的总积，如342即是50-54这一组限以下各组考生原始分数分布次数的总和；第4栏的数值，表示相对应组的原始分数在百分位量表上的位置，或该组原始分数的百分等级量数，如90.7就是45-49这组原始分数在百分位量表上的位置，亦即原始分数在45-49这一组限内的考生成绩在376名考生成绩中所占的地位，可解释为原始分数在45-49这一组限内的考生成绩高于90.7%的考生成绩，或者说，全体考生中有90.7%的考生成绩低于45-49这一组限内的考生成绩。

分组后的原始分数，可用下式求百分等级：

$$PR=\frac{\left[F_b+\frac{(X-L)f}{i}\right]}{N}\times 100 \qquad (7\cdot5)$$

(7·5)式中：PR为百分等级。

F_b为某原始分数所在组以下各组次数之和。

L为某原始分数所在组的下限。

f为某原始分数所在组的次数。

i为原始分数各组组距。

X为某原始分数。

N为原始分数分布的总次数。

现以表7－6某市376人英语考试的成绩为例，设某考生在该次考试中的原始分数为53分，求该生原始分数的百分等级。

由于我们已知某考生的原始分数（X）为53分，又自表8－6中得知该次英语考试的总次数（N）为376，53分所在组的下限

（L）为50，53分所在组的次数（f）为20，53分所在组以下各组的累积次数（F_b）为342，原始分数的组距（i）为5，便可直接将各数值用公式(7·5)求百分等级：

$$PR=\frac{\left[F_b+\frac{(X-L)f}{i}\right]}{N}\times 100$$

$$PR_{53}=\frac{\left[342+\frac{(53-50)\times 20}{5}\right]}{376}\times 100$$

$$=\frac{342+12}{376}\times 100$$

$$=94.2$$

通过计算得知，该生原始分数(53分)的百分等级为94.2，说明该生的英语考试成绩超过94.2%的考生，或该次英语考试中有94.2%的考生成绩低于53分。

未分组的原始分数，若要求百分等级，应先把考生原始分数按高低顺序排出名次，然后用公式(7·6)计算：

$$PR=100-\frac{100R-50}{N} \tag{7·6}$$

(7·6)式中：PR为某原始分数的百分等级。

R为某原始分数在考试总体中的顺序名次。

N为考试的总人数。

下面以某高校一次近代史考试成绩为例，计算未分组原始分数的百分等级。已知本次近代史考试共有160人参加，且按原始分数的高低排出了160人考试成绩的顺序名次，现用公式(7·6)求第84名考生原始分数（73分）的百分等级。

$$N=160 \qquad R=84$$

$$PR=100-\frac{100R-50}{N}$$

$$PR_{84}=100-\frac{100\times84-50}{160}$$

$$=100-52.2$$

$$=47.8$$

从计算结果可知，第84名考生的百分等级为47.8，表明第84名考生的73分在160人参加的近代史考试中，处于比52.2%的人差、或比47.8%的人优的位置。

百分等级常模有意义明确、易于理解、使用方便的优点。其缺点在于单位不相等，不同考试分数的百分等级不能相互比较，可靠度较之标准分数和T分数常模低（百分等级与标准分数的关

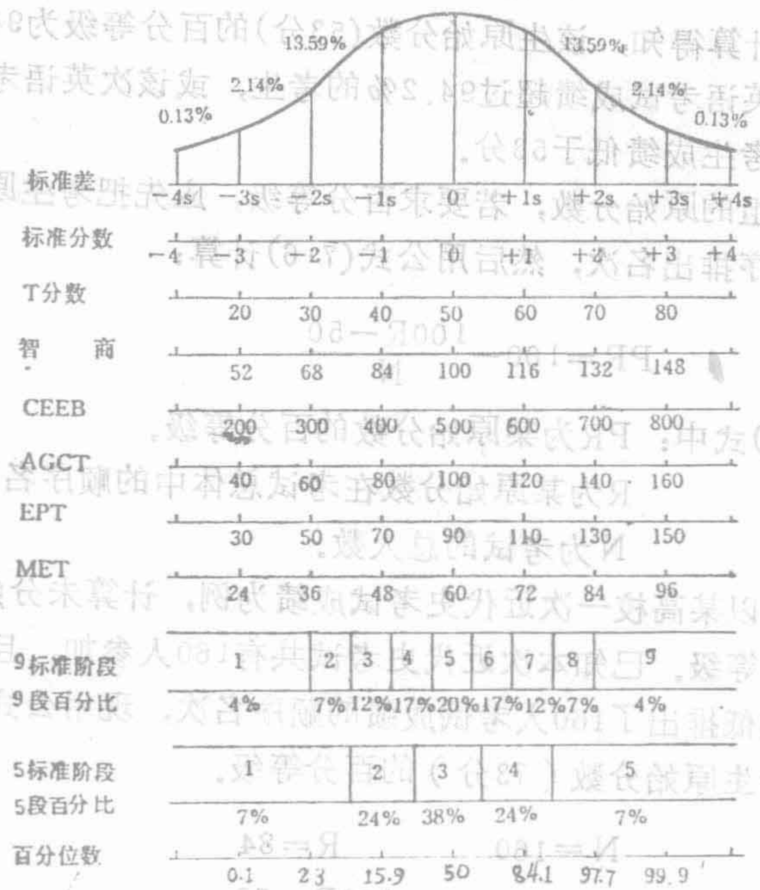


图 7－15　各种常模分数关系示意图

系，可参见附录表2）。

以上介绍的几种常模各具特点，适用范围及其效应也不尽相同。但是不管哪种常模，它都是一个转换考试分数的参照点和解释分数的依据，使考试分数具有某种实际意义。为便于对不同常模分数相互关系的了解，现将百分位分数、标准分数、T分数、5标准阶段分数、9标准阶段分数及其它几种常模分数的关系列于图7-15。

八、编写使用试卷的说明书

为了阐明考试设计者有关试卷适应范围、使用试卷的规则与要求，考试设计在完成了上述7个步骤的任务后，还需编写一份使用设计试卷的说明书（主观性考试无此步骤）对考试设计的依据，考试的目的、对象、内容与功能，印制试卷、答卷的规格与要求，使用试卷的原则和注意事项，测试时限，评分标准，常模及转换和解释分数的方法，以及其它必须明确的问题，分别作出简要的文字说明，便于使用试卷的单位能严格按照考试设计者的意图去实施，不致出现误用或滥用，以保证考试的质量。

完成了试卷使用说明书的编写，即是考试设计的结束，所编试卷就可正式用于考试。

第三节　考试实施

考试从开始设计到实施结束，本是一个有序的流程，各环节衔接紧密，不应相互割裂。但在考试的具体实践中，由于考试设计与实施主体的变化，即设计主体不一定是实施主体，一些社会性、区域性、地区性的考试尤其如此，通常是设计者与实施者各履其责、各司其职，还由于考试设计的超前与集中、考试实施的

滞后与分散等因素的限制，人们又把考试流程分为两大阶段，使考试设计与实施成为两个相对独立的流程，各有一套相应的程序。

从现行考试实施的流程看，基本可归为三类：一类是四步骤流程：编印试卷──→施测──→阅卷评分──→处理考试结果．各机关、学校、厂矿企业内部考试的实施流程，多属此类．第二类是六步骤流程：制订实施方案（计划）──→印制试卷──→设置试场──→施测──→阅卷评分──→处理考试结果．采用六步实施的考试，多为涉及面较宽、规模较大、具有一定社会影响的主观性考试（传统考试），如目前我国的毕业会考、预考、学科竞赛考试、合格考试、研究性考试、高考及各类选拔考试等．第三类是八步骤流程：制订实施方案──→印制试卷、答卷──→设置试场──→施测──→阅卷评分──→分数转换与统计分析──→编写考试分析报告──→处理考试结果．现今流行的常模参考性考试、目标参考性考试及其它客观性考试，均按此类流程实施．

现行考试实施采用的三类流程，在具体程序上差异至为明显，但都是多个环节的有机结合．就其科学性而言，八步骤流程要优于其它两类流程，故选取八步骤流程为例，分别介绍考试实施流程的基本步骤、内容与要求．

一、施测前各项工作的步骤与要求

在考试正式施测之前，主持考试实施的部门，应按三个步骤完成下列各事：

第一步，研读考试大纲和考试说明书，制订考试实施的具体方案．其主要任务是，根据考试大纲的要求及考试说明书规定的原则，针对考试实施所具备的各种条件，订出实施考试的整体工作计划，以及每一步骤的各项规程与细则，作为指导考试实施的总章程，以加强各实施环节的质量控制，使整个考试实施能朝着

预定的目标有序运行。因此，在考试实施的方案中，对考生报名、填报志愿、应考的有关事项，考试日期，工作日程，试卷、答卷印制的时间、地点、数量、格式、纸张与字体、字号规格，试卷的密封规则、运送方式、交接手续与保密措施，考场、试场设置的数量、地点、时间，各试场应考人数、设置要求，应考规则及违纪处理办法，监考人员的数量、条件、职责及工作细则，试卷管理人员的职责与纪律，评卷、联络、巡视人员的数量、条件、职责与工作细则，阅卷地点与时间，统计分析人员的职责与纪律，公布考试结果的方式、时间与方法，考生试卷、答卷原稿的处理措施，等等，应有细致周密的设计，并分项作出明文规定，提前以文件、文告或小册子的形式公诸于世，让有关人员熟知，便于实施过程中共同遵照执行，使考试实施的各个环节都有严格的质量控制。

第二步，核查被试人数，印制试卷、答卷和准考证。在这一步骤中，着重抓好三件好事：一是严格核实应考人数，做到准确无误，避免发生因试卷短缺而影响考试的正常进行，或由于试卷印制过多而造成浪费；二是抓好试卷、答卷的印制（若是用题库试题施考，还需精心拼配标准卷和复本）；三是抓好试卷、答卷的保管与发送。这三项工作，除核查被试人数一项可在主考单位内部进行外，其余两项都需与外部人员合作完成，责任性强，保密要求高，必须注重三条：其一，应有高度的保密措施。除对印刷厂要慎重选择外，试卷印制过程的每一环节都需有专人负责，并要选配可靠的监护人与之配合。考试所需试卷要一次印制完毕，不能中断或分次印刷。印制试卷的时间要科学安排，过早印刷，势必增大试卷保管与保密的压力，离考期太近，要确保试卷的印刷质量会有一定困难，若因偶然事件或特殊情况而造成逾期，还将殆误考试的实施。其二，把好试卷、答卷印制的质量关。试卷印刷的任务，应由设备好、技术力量雄厚的印刷厂承

担。试卷所用纸张、字体、字号的规格必须统一。当试卷、答卷的样稿印出后，需派专人或由试题编制者反复校对，确无差错、遗漏后，才能付印。印出的正式试卷，卷面墨色要均匀、字迹要清晰。为便于试卷的密封和保管，最好把试卷装订成册，将答卷分开印制，但试卷和答卷的式样、文字的印刷颜色可多种多样，以便同一试场的考生同时使用几种式样与颜色的试卷和答卷，作为防舞弊的措施之一。其三，在试卷印制完毕至施测之前这段时间，要注意四点：保管妥善，分发准确，密封规范，运送安全。具体要求是，印好的试卷要选择安全可靠、不易被他人察觉的场所存放，选派责任心强、确有看守能力的人看管。试卷的分发与包装，应以主考机关核准的人数为依据，并建立试卷分发记录薄，按照装封规则及分发的程序与顺序，依次进行试卷清点、装袋、密封、编号等工作，每个环节都要设核查员进行复查，确实无误，才能送入下一程序，最后由试卷分发、装袋、密封、编号、核查的负责人，在试卷分发记录薄上签字，以存档备查。至于各考区或考场试卷的运送，要根据运送的具体方式决定安全措施。若是各考区自行取卷，就要订出试卷领取与送交的程序和原则，为防止运送途中泄密或遗失，务必安排两三名精明强干的人负责押运。如果试卷是由主考单位自邮局寄往各考区，则应从机要系统以绝密类邮件寄出。在规定期限内，接收单位若未收到试卷或发现收到的试卷数量有出入，以及试卷袋破损、号码不符等问题，当立即电告试卷寄出单位，以便采取紧急补救措施。不管选用何种方式运送试卷，都必须提前一天将试卷送达各考场，保证施测按既定时间进行。

第三步，落实、训练监考和巡视人员，检查施测准备情况，设置试场。大规模考试的监考人员和巡视员，至迟在施测前一个月应落实，并给他们以必要的工作训练，使之明确监考、巡视工作的任务与要求，以便施测时正确履行职责。除此项任务外，该步

骤的中心工作是设置试场，任务包括打扫试场环境卫生，摆设试场桌凳，张贴试场示意图、试场规则、试场号、座位号，安排考生、监考人员考间休息场所等。试场设置应在施考前两天设置就绪。临考前一天，主考机关和巡视人员应对各试场，进行一次全面检查，看是否符合试场设置的总原则：既适合考生的应试心理、有利于应试者真实水平的发挥，又不致提供作弊的条件。是否达到规定要求：环境安静，通风条件好，试场光线适宜；试场四壁及应试桌凳上，没有与测施内容相关的文字或图表；座位前后左右距离恰当，座次随机编排，各试场的座位号与被试准考号一致；座次号码张贴位置同一；各试场的设置规格无明显差异；试场示意图清晰明了；试场编号无错漏等。与此同时，还要检查考前的其它工作的准备情况，如医务人员的配备，保卫工作的安排，施测所需物品的购置等。上述各项，若发现有不符要求的地方，当同时纠正。检查结束后，试场不得重新变动，除安排监考人员熟悉试场环境外，其他人严禁入内。

二、施测步骤的规则与要求

凡是正规的考试，其施测过程都要历经测前预备、施测、测试结束三个阶段。各阶段主试与被试的任务及要求如下：

（一）测前预备阶段

进入试场前，主试要明确考规考纪、工作职责及程序，熟悉试场设置情况、考间休息地点与测试的日程，备齐施测所需的各种材料，不得有短缺。按规定时间进入试场后，主试必须保持和悦、镇静的态度，不能给被试造成紧张气氛，并依照先后顺序完成正式测试前的各项任务：交待试场规则，分发试卷、答卷、草稿纸，对本试场考生与考号、座号核准，试前解说。指导填写卷头空白（如姓名、考号等），主试执行以上任务的基本要求是，试卷要当众启封，解说严格按统一的指导语进行，不允许随意增

减，解说时语言清楚，速度适中；试卷分发不能有遗漏；解说答卷方法时不能让考生翻开试卷；整个测前预备工作要在正式测试前结束，切忌占用考生应试时间。

（二）施测阶段

统一规定的测试时间一到，主试立即宣布：答卷开始！考生立刻翻卷解题。主试随即把测试起止时间写于黑板，没有黑板的试场，应重申一次测试的起止时间。在测试过程中，主试不得离开试场或作职责以外的事。当立（坐）于应试者之前，密切注视测试动态或在应试者周围缓步巡视，详细记录施测情况。对应试者有关解题方法不明确的提问，只能按统一规定给予答复，不能有任何正确作答的暗示。对测试中出现的违纪行为及突然事件，要按试场规则妥善处理，主试不得以任何借口拖延测试时间。应试者在测试期间，自入场到出场均以试场规则为准则，凡行为不符规则要求者，当视其情节轻重，依照考规考纪进行处理。

（三）测试结束阶段

到了规定的测试终止时限，主试立即宣布：停止答卷！接着收试卷及草稿纸，若是试卷与答卷分开的考试，则应先收答卷，再收试卷，铅笔、橡皮等物应最后收。试卷收齐后，按本试场考号序次进行清点，理顺每份试卷的页码，并与实际参考人数核准。若无遗漏，随后处理试卷：首先，对监考人员使用、多余、缺考、未答、舞弊的试卷，分别在卷头或总分处写上“作废”、“缺考”、“白卷”、“违纪”字样，再与其它试卷一起装入试卷袋，填写内封与封筒，按统一规格密封，然后退出试场。

施测是考试实施过程中的核心步骤，也是影响考试成败的关键环节，其质量应有严格的控制措施。在统一的测试时限内，每一试场的主试和被试，都务必执行同样的指令，说相同的话，作相同的事，在相同的时限内完成相同的任务。

三、施测后各项工作的步骤与要求

从考试实施的全部工作讲，结束施测，可算完成了一半任务。在施测后期，还有阅卷评分、质量分析、编写考试报告、处理考试结果四大任务。按照八步骤实施流程，当分四步进行。为叙述的简便，现根据任务的性质归并为两大步骤。

（一）阅卷计分与质量分析

这一步骤的主要任务是：组织阅卷队伍，培训评卷人员，准备阅卷所需物品及计算工具，阅卷评分、合分、登分、计分和分析考试质量等。评分客观、登分无误、计分准确、分析可靠，是对阅卷计分和质量分析工作的基本要求。

至于阅卷计分与质量分析的具体工作环节，一般取决于阅卷评分的方式、方法，而试卷评阅和质量分析的方式与方法，又是因考试的性质而变化的。因此，现行的两大类考试（主观性考试与客观性考试）在试卷评阅和质量分析的方式、方法方面，区别至为明显，有分别讨论的必要。

1．主观性考试评阅试卷与质量分析的方式及方法步骤。

由于主观性考试的试题绝大多数是论文式题目，试卷评阅与质量分析的方式在客观上受到限制，所以它很难采用光电阅读、电子运算等现代技术手段完成上述任务，目前还只能是打“人海战术”，即阅卷、评分、合分、登分、计分以手工方式为主，质量分析以定性分析方法为主。

主观性考试阅卷评分与质量分析的具体工作环节是：

（1）建立临时评卷机构。试卷的评阅是一项时间紧迫的工作，一般要求在施测后半月内结束。主观性考试因全靠手工劳动，考生与评卷人员的比例通常是20：1，一些规模宏大的考试，往往要动用大量的人力和物力，象我国的高校入学考试，每年需抽调近10万人参与此项工作，即使一个省也是数千人的评卷

队伍，这就必须建立一个临时的评卷机构，负责试卷评阅的组织与管理。

临时评卷机构由主考机关考务负责人牵头，并设立中心联络组，再根据工作的实际需要，分三级设立若干个小组：第一级是学科领导小组，如语文、数学、外语评卷组等；第二级是职能小组，如评卷指导小组、试卷管理组、合分组、登分组、质量分析组、后勤组等；第三级是阅卷组。其组成形式有两种：一种是以试题划分，一个组负责评阅一道或几道大题，素称分题阅卷法；另一种是按人员分组，每组承担多少份试卷的评阅，所以叫分组

表7－7 考试评卷的组织机构及其职责

组别	工作职责
学科领导组	全面负责学科试卷的评阅和质量分析。
评卷指导小组	培训评卷人员；根据标准答案制定评分细则；解决阅卷、评分中的疑难或争议性问题；把握评分标准及评阅质量。
阅卷组	判别答题正误，评定各题分数。
合分组	根据阅卷组评定的各题分数，计算各考生全卷总分。
登分组	按考区和考号将每个学生所得分数登记造册。
质量分析组	对学科考试作抽样分析，评价试题及考试质量。
试卷管理组	对外负责试卷的领取与运送，对内负责试卷的分发回收与保管，以及评卷进程的统计与公布。
后勤组	设置评卷场所、购置评卷物品，负责评卷人员的食宿及文娱活动的安排。
联络组	向各组传达领导组的有关精神，了解评卷进度及出现的问题，沟通信息，协助处理有关问题。

承包阅卷法。各组的工作职责如表7－7所示。

表中各组所需的工作成员，除临时评卷机构负责人由主考机关专司考务管理的领导人充任、部分联络员由考机关工作人员担任外，其他成员，均由主考机关从社会有关部门临时雇请，并指定学科评卷组负责人。至于各职能组的负责人，则由学科评卷领导小组确定后，再报请主考机关核准。

（2）培训队伍。培训评卷人员，是主观性考试阅卷评分过程中的一个重要环节，它直接关系到能否有效控制评卷质量的问题。其原因在于两个不稳定：一是试题标准不稳定。主观性考试多在临考前期命题，没有统一的考试大纲和编题计划作为同类考试各次命题的依据，又无检验试题质量的预测环节，各次考试的试题长度、难易标准、题型类别等不稳定，阅卷原则和定分标准一次一个样。所以，每次评卷之前，都有一个再次了解、重新定调的问题。二是评阅队伍不稳定。主观性考试的评卷人员不仅数量大、涉及面宽，而且每次都得临时雇请、重新组合，虽然每次有老成员参与，但毕竟是少数，其中大部分成员是新手，对他们必须进行试卷评阅基本技能的训练，才能负起阅卷评分的职责。

为提高培养的效率，培训工作分两步进行为宜。首先集中各职能组和阅卷组组长，培养阅卷的骨干力量，让他们在了解评卷原则与要求、熟悉标准答案的基础上，进行抽样试评，以掌握评分标准，并针对当次考试的实际情况定出执行评分标准的细则，为整个试卷评阅提供具体依据及参照样本（经过试评选出的标准评分样卷）。然后再培训全体评卷人员。

（3）阅卷评分、合分、登分与质量分析。在主观性考试的评卷过程中，这四项工作既是严而有序的，同时又是四位一体的。一方面，评卷应按阅卷评分—→合分—→登分—→质量分析的序次进行，前一步骤不提供素材，后一步骤就没有开展工作的条件，故不可程序颠倒。另一方面，四项工作虽各有职责，但都

是评卷这一中心任务的合理分工，各项工作开始有早晚，又必须在大致相同的时限内结束。因此，主观性考试的评卷工作不象客观性考试那样，完全按照设计程序直线推进，它是几方面的工作同步进行，需有各组的密切配合，才能使整个评卷工作协调运转，进而实现对评卷质量的有效控制。

（4）综合各科成绩，计算考生总分。计算总分是否作为阅卷计分工作的独立环节，当视考试的具体情况决定。若是单科考试，该项任务可由登分组承担，不必与阅卷评分环节分开。多科性的综合考试，则应独立于其它环节，由临时组建的计算站负责各科成绩的累计，以及考生成绩报告单或成绩手册的编印。同时，也可委托专业计算中心负责运算。

2．客观性考试评阅试卷与质量分析的方式及方法步骤。

在阅卷评分与质量分析的任务与要求方面，客观性考试与主观性考试并无本质区别。但因客观性考试具有五大特点：全部采用客观性试题，或以客观性试题为主体，辅之以少量非客观性试题；各次考试的题型、试卷长度与标准基本相同；试卷与答卷分离；评卷队伍较为稳定，而且数量小、水平比较均衡；不以计算出各考生的原始总分作为衡量考生成绩优劣的最终依据。所以，客观性考试评阅试卷与质量分析的方式、方法和具体工作程序，均不同于传统考试：在阅卷、评分、合分、登分、计分及分数转换的方式上是以机器为主，而不象主观性考试主要采用手工方式；在质量分析的方法上，实行定量分析与定性分析有机结合，即在量化分析的基础上进行定性分析；在评卷工作的具体程序上，是纵向推进，四步一线，互不躐等，严格按照设计程序阅卷评分（包括合分、登分、计分）、分数转换、质量分析的序次进行。

以上四个环节的工作，主要靠光学阅读器和电子计算机。两种试题（客观性试题与非客观性试题）并用的客观性考试，其阅卷评分则采用综合手段，考生的原始总分由几种彼此独立而又相

互走向一致的分数合成. 基本作法是: 客观性试题的评阅借助电脑技术完成. 至于非客观性试题的阅卷评分, 现今世界各国并无统一的方法, 使用较普遍的是两种: 一种是沿用传统考试的阅卷评分办法, 或由资深考官打印象分, 或由评卷人员依照标准答案和评分细则判卷定分. 另一种是标准阶段评分法, 亦即分解评分法, 通常是根据正态分布原则, 把考生总体成绩的分布划为5段、7段或9段, 再按各段在正态曲线下所占面积的比例评定分数或成绩等次. 采用这两种方法评出的原始分数, 要与同一试卷客观性试题部分的分数进行比较, 检验每个考生两种试题的得分有无线性关系, 走向是否一致. 倘若同一试卷客观性试题得分低, 而非客观性试题得分很高, 还应分析具体原因, 并作必要调整, 才能将两种分数合成考生的原始总分. 最后通过常模转换, 把原始分数变为等值的标准分数（导出分数或整理分数）作为考试成绩的最终结果.

（二）处理考试结果

无论是主观性考试还是客观性考试, 处理考试结果都是考试实施的终局环节. 但在处理考试结果的方式、步骤和具体任务方面, 两种考试又互有区别. 现比较如下:

首先, 主观性考试在算出考生原始总分之后, 即可转入考试结果的处理. 客观性考试算出考生原始总分后, 还要历经分数转换、质量分析和数据资料综合整理三个环节, 才进行考试结果的处理工作.

其次, 客观性考试在正式处理考试结果之前, 必须综合整理各种量化数据, 并以分析报告的形式, 对整个考试的难易程度、区分能力、信度、效度、考试分数的分布情况、误差率等, 分门别类地列出量化指标, 给予明确的解释和科学的评价, 作为正确处理和使用考试结果的依据. 主观性考试则无此项任务.

最后, 主观性考试在处理考试结果的过程中, 给考试结果使

用单位只提供考生个人的档案和考试分数，给考生本人只报告考试的原始分数。客观性考试在处理考试结果时，给考试结果使用单位或个人提供的依据，则不单是考试分数，还附有反映整个考试质量的全部统计数据和完整的分析资料，使考试结果使用者能全面了解考试质量，以及各考生成绩在考试总体成绩中的地位，对每个考生掌握知识的情况和已经具有的能力水平能作客观判断，做到选才准确、用才合理，有利于提高考试的社会效应。

第八章

考试的数据整理与质量分析

科学的数据整理和质量分析，是提高考试质量、正确使用考试结果、有效发挥考试功能的重要前提。

人类社会所进行的一切测量，都是按照一定的法则给被测对象指派数字。考试是对人的德、学、才、识、体个别差异的测量，主试给被试的测试结果评定分数，也是依据一定的法则给被试的测试结果指派数字。这种数字能在多大程度上反映两种质与量（被试受测方面客观存在的质与量；考试实际测得的质与量）的一致性，则取决于指派数字的法则，即对被试测度、甄别和评价测得结果的方法与手段。从理论角度讲，凡设计标准、程序科学、实施规范的考试，它所测得的质和量与主试欲测的质和量及被试受测方面客观具有的质和量，应是三元合一的，即完全吻合、毫无误差。但在实际考试中，三方面的质与量不可能完全合一。因为任何一种考试的结果，都无不含有两种成份：欲测并已测得的质与量，与测试目的无关因素所造成的误差。也就是说，“凡测量都有误差”，考试亦是一种测量，测试结果中的误差客观存在，只是由于各种考试对导致误差因素的可控能力不一样，而在测试结果中误差所占的比重有大小之别。所以，平时所说的考试质量高低，实际指的是考试结果中误差比重的大小。所谓分

析考试质量，也就是运用统计方法检查考试结果中的误差比率，或者说，检查考试结果反映被试真实水平的程度。

现今世界各国在考试质量的分析，以及对各种统计量数之可靠性的检验方面，途径各殊、方法多样、内容不一，若取其共同、择其要者，也可归结为：常规分析与理论分析两种方法；定性、定量及定性与定量相结合三种手段；转换分数、整理数据、质量分析、可靠性检验四个步骤。现依次探讨于下：

第一节 转换分数

一、转换分数的概念、目的与意义

转换分数与分数转换属同一概念。其含义是把值不等、义有别的原始分数转换成某种意义相同、数值相等的分数，实质是使考试分数标准化的一项技术措施，也可以说，是对变不等值分数为等值分数所用方法的一种称谓。由原始分数转换而成的等值分数，素称整理分数或导出分数，即是"用数字所表示的常模"。

在计算出考生的原始分数后，采用统计方法对原始分数进行转换的目的，是要把考试分数由不等值变为等值，使每个考生不同科目的分数能相加，以便组合成一次考试的总成绩，并与同次考试中其他考生的成绩可比，以求准确判定每个考生的成绩在同一考试总体中所占的位置。

至于进行考试分数转换的意义，则是显而易见的。人所共知：考试的原始分数对一次考试来说，它只是一组含义不清的数据，并不直接表明一次考试质量的高低；而对考生个人来说，其所得原始分数，也仅仅是测量量表上一个孤立的读数，它本身不具有任何实际意义，既不表明成绩优劣，也不能比较出其它考试

成绩的好坏。比如某考生的历史成绩为86分，外语考试为50分，我们就无法判断该生两种成绩的好坏，也许50分在同次外语考试中属于上等成绩，而在同次历史考试中，86分又可能属于下等成绩，也许该生的外语和历史成绩都处于领先地位。同时，我们也不能把86分与50分相加，作为该生外语、历史两科考试的总成绩。因为，要使考试分数能互比或相加，就必须有相同的意义、相同的单位和反映分数的共同参照点。考生所得原始分数，虽然有相同的计量单位（以1分为单位）和共同的参照点（以0分为起点），但因不同科目考试的份量、难易度、评分标准，以及考试结果中误差比重的大小不同。所以，即或都以0分为起点，可零点的实际水平有别，都是以1分为单位计分，1分的值却不等。这种计量单位和参照点，对不同科目的考试，以及同一科目不同时间举行的考试来说，并不具有相同的意义，没有互比或相加的条件，进而也就无法对考试分数作出有确定意义的解释。正由于此，欲使考试的原始分数具有实际价值，就必须按照一定的程序和方法，把原始分数转换成具有统一意义、相同单位和共同参照点的某一量表上的等值分数。

二、转换分数的方法

转换原始分数的基本方法是：在试卷容量、难易程度、评分标准大致相同，考生分数的平均值和标准差与常模分数基本一致的前提下，将考生所得原始分数以常模为统一的参照点进行转换，使全体考生的原始分数为同一量表上具有相同单位和意义的值。若是多科性考试，在分别整理了原始分数之后，需进行加权处理，然后再将各科分数合成，作为一次考试的总成绩。

经过转换的分数称为导出分数。这种分数既可把某考生的分数与其他考生的分数比较，也可与该生自己不同科目或不同考试的分数比较。导出分数的种类很多，如标准分数、T分数、百分位

数，等等。转换原始分数的公式，根据转换分数时所参照的常模类别确定，比如，参照的是标准分数常模，就应使用标准分数的计算公式：$Z=\frac{X-\overline{X}}{S}$若参照的是T分数常模，其转换公式则为：$T=10\frac{(X-\overline{X})}{S}+50$，或$T=10Z+50$。有关此类转换公式及具体计算方法，已在第七章第二节关于建立常模部分作了较为细致的讨论，此不重述。

第二节　整理考试数据

一、整理考试数据的概念与意义

一次考试结束后，人们根据考生答题结果评出的原始分数，便是主试最先得到的反映考试质量的数据。它既是每个考生考试成绩的量化，又是整个考试结果的量化，也是检验考试质量优劣的原始依据。但这些原始数据尚处于不具条理性的无规律状态，除能看出单个数据的大小外，无法从中看出一次考试情况的整体面貌，以及各个考试分数的意义与次数分配的情形。因此需按照适当的法则使之条理化，理顺数据的相互关系，使蕴藏在大量原始数据中的反映考试质与量的真实情报实现外化，以便人们准确了解和掌握，并形成一次考试分数分布的完整直观概念。这种对原始数据进行科学的分组或归类的过程，就叫数据整理。

数据整理的具体任务，是制作考试分数的分布图表，借以反映一次考试全体考生所得分数散布于不同分数段的状况，亦即全体考生所得分数在各个分数段（或区间）次数（或频数）分配的特点。因之，人们通常将数据整理的结果表述为分数分布、次数

分配或频数分布。按一定法则将大量原始数据归入图表后，一次考试分数的整体状况便一目了然。诸如，考试的最高分与最低分各是多少？整个分数的次数分配是集中于高分段、低分段，还是中间层？考试分数的全距是多少？考试的内容是偏难、偏易、还是适度？分数的集中趋势与离中趋势如何？等等，都可从图表中得到清晰的直观了解。所以，经过数据整理，以图表形式揭示一次考试分数分布的情形，是直观了解考试整体状况的基本措施，也是进行考试质量分析的基础工程。

二、整理考试数据的方法

在整理考试数据的实际工作中，人们常常是以图表形式反映一次考试分数的整体分布状况，采取求集中量数和离中差数的办法，描述考试分数的集中趋势与离散程度。其基本方法是：

（一）制作考试分数分布图表

随着现代统计方法的发展，中外考试界用于反映考试分数分布状况的图表种类日益增多，现以次数分配表、多边图、散点图、正态分布图四种常用的图表为例，分别陈述其制作的方法和步骤。

1．次数分配表。

它是以表格形式按分数段反映一次考试分数的次数分配特点。传统考试及现行的主观性考试，一般都采用此种方法描述考试分数的分布。下面以一次72人参加的历史考试分数为例，说明次数分配表的制作方法。72名考生历史考试的成绩如表8-1所示。

表8－1　72名考生历史考试分数

81	73	64	59	94	82	67	75	90
48	57	86	75	93	88	72	65	77
60	79	70	81	56	64	73	99	87
66	50	45	80	82	67	74	68	92
54	57	87	64	76	55	61	65	70
89	71	49	60	72	70	83	67	74
90	68	75	65	56	59	73	81	78
70	75	62	69	84	53	56	71	69

若要对这一组无规律的原始数据进行整理，首先需按下列步骤进行分组：

（1）求全距。求全距即是找出一次考试分数的最高分数（最大值）与最低分数（最小值）之间的差数。全距又称两极差，在数理统计中用R表示，其计算公式为：

$$R = X_H - X_1 \quad (8\cdot1)$$

（8·1）式中：R为全距（两极差）。

X_H为一次考试的最高分数。

X_1为一次考试的最低分数。

现已知72名考生在历史考试中的最高分数为99分，最低分数为45分，用公式（8·1）求其全距：

$$R = X_H - X_1 = 99 - 45 = 54$$

结果表明，该次历史考试分数的全距为54。

（2）确定组距与组数。所谓组距，即是每一组所含分数中最大数与最小数之间的距离。组数表示一次考试的全部分数按组距归类后小组的个数。确定组距是为了分组，分组的目的在于缩简运算程序。

组距的大小与组数的多少，应根据各次考试数据的实际情况决定，以方便统计为原则。一般情况是，将一次考试的分数分为10～20个小组。出于确定组距中点的简便，大都用单数作组距，如0—3，0—5，0—9等，其中又以取5为组距的最为普遍。求组距的方法为：全距加1除以组数。因此，当一次考试分数的全距求出后，就可确定组数，再用全距加1除以组数，所得商数即可作为该次考试分数的组距。例如，我们自表8-1得知72名考生历史考试分数的全距为54，若决定将全部考试分数分为11个组，其组距则为$\frac{54+1}{11}=5$，我们就可用5作该次考试分数的组距。

（3）决定中点与组限。组限是指分组后每一组所含数的两个端点（起点与终点）的数值，这两个数值表明每一组的界限，最小一端的数值谓之下限，最大一端的数值谓之上限。如45—49.9一组中，45为下限，49.9为上限。决定组限也就是定出一次考试分数的具体组距。例如我们已将72名考生历史考试分数的组距定为5，本次历史考试的最低分为45，就可以45为最小一组的下限，依次排除该次历史考试分数的组距：45—49.9，50—54.9，55—59.9，60—64.9，65—69.9，70—74.9，75—79.9，80—84.9，85—89.9，90—94.9，95—99.9。从组距中可看出，在最小组中应包括一次考试的最小分数，而在最大一组中应包括一次考试的最大分数。

一次考试分数的具体组距与组限确定后，继而决定组距的中点，以此作为每组各个分数的代表数值。组距中点的计算公式为：

$$K=L+\frac{U-L}{2} \qquad (8\cdot2)$$

（8·2）式中：K为组距中点。

L为组距的下限。

U为组距的上限。

例如在45—49.9一组中，上限（U）为49.9，下限（L）为45，代入公式（8·2），该组的中点则为：

$$K=L+\frac{U-L}{2}=45+\frac{4.9}{2}=47.45$$

为计算方便，组距的中点数值（组中值）一般保留一位小数。因此，45—49.9一组的组中值可按47.5计算。

（4） 制作次数分配表。一次考试分数的全距、组距及其组中值一旦确立，即可着手次数分配表的绘制。这一步骤包括计算次数（数频数，划记）、分组归类、记次数三项任务。次数分配表的制作需按三步进行。首先在次数分配表的组距（i）和组中值（K）两栏内，分别写下组距和各组的组中值。再把各个考试分数按其所属组限范围分别归入各组。其方法是根据考试分数的自然序次逐个往下看，将每个分数归入所属组，每归组一个分数，就用划线或写正字的办法，在所归组相对应的登记栏内划一直线或给正字加一笔，最后计算出各组相对应登记栏中的直线或正字笔画的总数，并以数字写入各组相对应的次数（f）栏内。这样，一次考试分数的次数分配表就算绘成了。如表8-2所示。

2.散点分布图。

它是一种最简易的分数次数分配图，只适合于小规模考试的数据整理（考生在200人以下的考试）。散点分布图是以一横坐标上的刻度代表一次考试大小不同的分数，用小圆点表示每一分数在同次考试中出现的次数，如图8-1所示。

制作散点分布图的方法是：先画一横轴，并按等距原则标明刻度。若实行的百分制，就标100个刻度（0—100），5分制即标5个刻度再标明刻度数字，就成了一个横坐标。若是100分制，

横坐标的起点为0，终点数值为100。也可用一次考试分数最低分作为横坐标的起点数值，用最高分作终点数值。然后按照考试分数的自然序次一个一个往下读，每读一个考试分数，就在横坐标刻度的相对应处点上一个小圆点。这样，读完全部考试分数，散点分布图也就制作完毕。

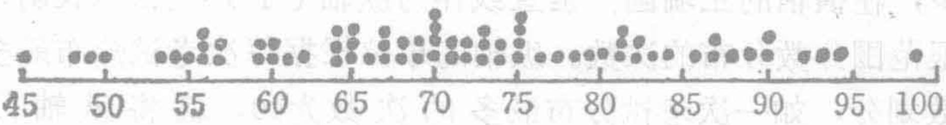


图8－1 72名考生历史考试分数散点分布图

表8－2 72名考生历史考试分数次数分配表

组 距（i）	组中值（k）	登 记	次数（f）
95—99.9	97.5	正	1
90—94.9	92.5	正	5
85—89.9	87.5	正	5
80—84.9	82.5	正 正	8
75—79.9	77.5	正 下	8
70—74.9	72.5	正 正 正	13
65—69.9	67.5	正 正 一	11
60—64.9	62.5	正 下	7
55—59.9	57.5	正 下	8
50—54.9	52.5	下	3
45—49.9	47.5	下	3
Σ			N＝72

3．次数分布多边图。

制作次数分布多边图分四步进行：第一步，在方格绘图纸上作一横轴（X），以表示一次考试分数的组距，一次考试分数决定分多少组，就将横轴分为多少段。每一段的单位按组距的单位确定，若组距是7个单位，横轴的每一段就定为7个单位。第二步，在横轴的左端画一垂直线作为纵轴（Y），用以表明每一组限范围分数分布的次数。纵轴的单位根据每次考试分布最多的次数划分，如一次考试分布最多的次数为40，就将纵轴分为40个单位。第三步，在各组组中值与各组次数两直线相交处作次数标记（圆点或圆圈均可）如图8－2所示。第四步，把表示各组分数分布次数的标记用直线连接起来，形成反映分数分布整体状态的次数分布多边图。现以表8－2内的次数分布为例，绘制72名考生历史考试分数的次数分布多边图。

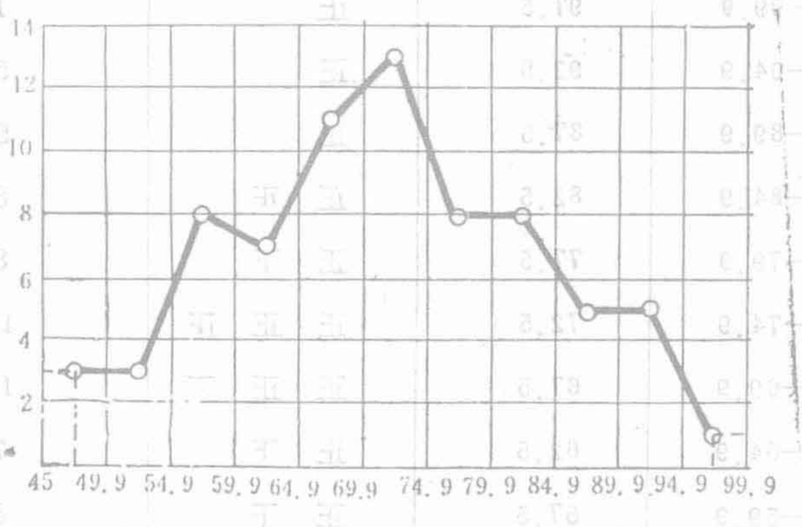


图8－2　72名考生历史考试分数次数分布多边图

4．正态分布图。

若以正态分布曲线反映一次考试成绩的总体分布特点，其分数的分布曲线可分为正态、正偏态、负偏态三种形式，如图8－3所示。

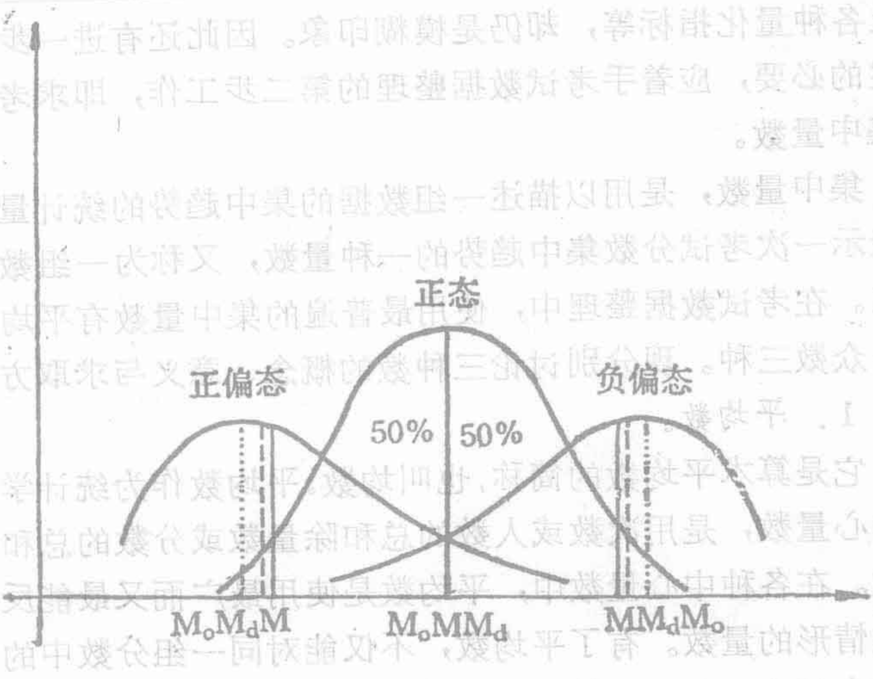


图8－3　分数分布的三种曲线形态

经过对考试分数的整理，如果中数（M_o）、众数（Md）与平均数（M）相同，那么整个考试分数的分布曲线将呈正态，中数、众数将重合于平均数（同时落在对称轴上），说明考试分数的次数分配在对称轴两边」占50%，同时表明考试内容难易适度，试题的鉴别能力强，考生成绩主要集中于中间段，高低两端得分者较少。倘若考试分数的平均数小于众数和中数，整个考试分数的分布曲线将呈负偏态，表明考试内容偏易，绝大多数考生的成绩居于高分段。如若考试分数的平均数大于众数和中数，整个考试分数的分布曲线便成正偏态，说明考试内容难度太大，中下成绩的考生占了考生总体的大多数。

（二）求集中量数

经过绘制考试分数次数分配图表，使考试所获原始数据结束了杂乱无章的状态，成为一组有条理的数据。通过次数分配图表，虽可看出一次考试的总体分布情形，但对次数分配的具体概念、含

义及各种量化指标等，却仍是模糊印象。因此还有进一步度量和描述的必要，应着手考试数据整理的第二步工作，即求考试分数的集中量数。

集中量数，是用以描述一组数据的集中趋势的统计量，也就是表示一次考试分数集中趋势的一种量数，又称为一组数据的代表值。在考试数据整理中，使用最普遍的集中量数有平均数、中数、众数三种。现分别讨论三种数的概念、意义与求取方法。

1．平均数。

它是算术平均数的简称，也叫均数。平均数作为统计学上的一种中心量数，是用次数或人数的总和除量数或分数的总和所得的商数。在各种中心量数中，平均数是使用最广而又最能反映一组分数情形的量数。有了平均数，不仅能对同一组分数中的每一分数作出评价，而且可将一组分数与其它各组分数进行比较（在分数等值的情况下）。求取平均数的公式为：

$$M=\frac{\Sigma X}{N} \tag{8·3}$$

（8·3）式中：M为平均数。

Σ为各数相加求和。

X为每一量数（分数）

N为总次数或总人数。

2．中数。

中数又称中位数，属中心量数的一种，常用Md表示，有时也用Mdn代表。中数在计算和意义上都与平均数不同，平均数是一次考试分数的平均值，反映一个考试总体的平均水平，中数是处于一组数据之中间位置的那个数据，一般不能代表一次考试的平均水平。只有在考试分数的次数成标准正态分布的情况下，中数才与平均数具有同等意义，平均数的计算是用考试的总分数除以考试的总人数（或分数分配的总次数），中数是以一次考试分

数的分配次数为依据，将全部考试分数按数值大小依次排列，然后由大小两端的任意一端数起，数到总次数一半的那个分数，就是中数。因而，在一个量数系列中，中数上下的次数恰好各占一半。计算中数的方法有三种：

（1）在考试分数分配次数的总和为奇数的情况下，其计算公式为：

$$M_d = \frac{N+1}{2} \quad (8\cdot4)$$

（2）若考试分数的分配次数总和为偶数，则先找出处于中点数的两个数，再相加除以2。以10名考生的政治考试分数为例，现已知10名考生的政治考试分数为：68，72，84，65，93，68，59，73，67，81，再按下列程序计算中数：首先排列分数的序次：59，65，67，68，68，72，73，81，84，93。再从大小两端数分配次数，当大小两端分别数到总次数一半的那个数时，就分别将那两个数抽出，然后相加除以2，即得中数。如上列10个分数，从低分一端数至总次数一半的那个数为68，而从高分一端数至总次数一半的那个数为72，两数相加除以2等于70（（68＋72）÷2＝140÷2＝70）亦即中数。

（3）若考试分数已制出次数分配表，则按公式（8·5）计算：

$$M_d = L + \frac{\frac{N}{2} - F_2}{f} \times i \text{或} M_d = U - \frac{\frac{N}{2} - F_1}{f} \times i \quad (8\cdot5)$$

（8·5）式中：M_d为中数。

L为中数所在组下限。

N为总次数。

F_1为中数所在组以上各组次数之和。

F_2为中数所在组以下各组次数之和。

U为中数所在组之上限。

f为中数所在组的次数。

i为组距。

3．众数。

众数也是一种中心量数，又称为范数，通常用M_o表示。众数有近似众数、理论众数和观察众数之分。考试分数中的众数，就是一次考试的分数中出现次数最多的那个分数。如果已将分数整理出次数分配表，次数最多的那个组的组中值，即为近似众数。在考试分数的次数对称分布的情况下，众数、中数和平均数相互重合。若次数分布不对称，则三个中心量数互不相同，一般情况是M与M_o的距离约等于M与M_d的距离的三倍，以公式表示则为$M-M_o=3(M-M_d)$。所以，近似众数可用公式$M_o=M-3(M-M_d)$求得。如果按考试分数的次数分配表列出的次数计算众数，则用（8·6）求得：

$$M_o=L+\frac{f_2}{f_1+f_2}\times i \text{或} M_o=U-\frac{f_1}{f_1+f_2}\times i \qquad (8\cdot6)$$

（8·6）式中：M_o为众数。

L为众数所在组的下限。

U为众数所在组的上限。

i为组距。

f_1为众数所在组下一组的次数。

f_2为众数所在组上一组的次数。

集中量数主要描述一次考试的集中趋势，即反映一次考试总体成绩的优劣，或全体考生的平均水平如何。因而，仅有平均数、中数、众数等集中量数，还不能对一次考试成绩作出全面、细致的描述，尚须求差异量数。

（三）求差异量数

差异量数是表示一组数据散布状况或离中趋势的量数，故又

称差异量数为离中差数。差异量数既适用不同单位资料差异的比较，也适用资料单位相同而平均数差别很大的资料的比较。差异量数的大小，反映各数值分布范围的大小及变动情况，表明各数据间相互差异或一组数据离散的程度。在考试分数的整理中，主要用于各考生成绩之差异情形的描述。整理考试数据使用最多的差异量数，除全距外，还有如下几种：

1．标准差。

标准差是数理统计中最可靠、最合理、使用最普遍的一种离中量数。标准差常以S、SD或希腊字母σ表示。若以S代表标准差，X_1^2 X_2^2 X_3^2……X_n^2代表一组数据内每个量数与平均数的代表差平方，N代表总次数，则可用公式（8·7）计算标准差：

$$S=\sqrt{\frac{X_1^2+X_2^2+X_3^2+\cdots\cdots X_n^2}{N}}=\sqrt{\frac{\sum X^2}{N}} \quad (8\cdot 7)$$

可见，标准差是一组数据内的每个量数与算术平均数的代表差的平方的均数之平方根。由于在求标准差时，先使各代表差自乘消除了负数，然后才求算术平均数，最后求均数的平方根，所以它在计算上补救了平均差的缺陷。再者，因求标准差一律用平均数作中心量数，而不用中数或众数，故又将标准差称为均方差，但标准差的概念较为确切。

2．平均差。

平均差系离中量数之一，它是一组数据内的每一个量数与中心量数（用平均数或中数均可）离差的平均数，多用AD和MD表示。平均差用公式（8·8）求得：

$$MD=\frac{\sum|(X-\overline{X})|}{N} \quad 或MD=\frac{\sum|x|}{N}$$

$$或MD=\frac{\sum d}{N} \quad (8\cdot 8)$$

（8·8）式中：MD为平均差。

$|\ |$为绝对值。

X为某量数。

$\overline{X}$为某一组量数的平均值。

x为某一量数与平均值的离差

$(x = X - \overline{X})$

d为某一量数与平均数的差数（不计正负号）。

Σ各差数相加求和。

N为总次数或总人数。

平均差在计算时，各量数与均数的差数一律不计正负号，全以正数对待。否则，离差的总和就等于零，进而也就无法求平均差。这是以平均差作离中量数的最大不足。

3．四分差。

四分差亦为离中量数之一，它是一组数据内各量数与中数相对应的差异量数。四分差用Q代表，通用而简便的计算公式为：

$$Q = \frac{Q_3 - Q_1}{2} \qquad (8\cdot9)$$

（8·9）式中：Q为四分差。

Q_3为第三个四分位数。

Q_1为第一个四分位数。

在整理考试数据时，求四分差的基本步骤是：先将一次考试分数由低分到高分依次排列，再把考试分数分配的总次数分为等量的四段，也就是确立三个四分点：Q_1，Q_2，Q_3。若从最低分一端算起，一次考试分数分配总次数的25％处便是第一个四分点（Q_1），总次数的50％处为第二个四分点（Q_2），亦即中数，总次数的75％处为第三个四分点（Q_3）。若一次考试分数是标准的常态分布，Q_2这个四分点的上下两端的次数均等，各占考试分数分配总次数的50％。在这种情况下，Q_2上下两端

中任何一个分数与Q_2的离差，都将小于四分差。

集中量数和差异量数是统计学中两个相互对应的量数，前者表示一组数据的集中程度，后者表示一组数据的离中或离散程度；集中量数描述数据的集中趋势，差异量数描述数据的离中或离散趋势。一次考试分数的差异量数越小，说明考生之间的成绩越接近，差异量数越大，则表明各考生之间的水平差异明显。

第三节　考试质量分析

分析考试质量是一项复杂而细致的工作，目前，一般是借助电脑按照一套依靠数理统计方法建立的程序，对试题及考试结果进行难度、区分度，信度、效度和诸种因素的分析，以获得反映试题及考试质量指标的各种数据。对于这一问题，由于我们将在《试题功能与考试质量分析理论》一书中作专门探讨，此处只阐述考试质量的常规分析法，即“四度”分析法。另外，因篇幅所限，有关“四度”计算方法只给公式，不举例。

一、难度分析

在各种考试中，无论使用哪种类型的试题，一般都用P表示试题的难度值。但是，不同类型试题之P值概念不尽一样。如选择题的P值使用的是通过率的概念，而问答题或论述题的P值，使用的则是答对率的概念。因此，所谓难度，就是考试试题对应试者实际水平的适合程度，亦即一次考试中答对或通过某道试题的考生数在其总体中所占的比例。试题的难度指数最高为1，最底为0，试题难度在0.35～0.8之间为宜，全卷的平均难度多控制在0.5～0.6之间。难度过大或过小，对试题的区分能力及考试结果的信、效度均有影响。

试题的难度可用多种方法计算，现选举5种不同的计算公式。

（一）问答题难度公式

$$P=\frac{\bar{X}}{a} \qquad (8\cdot10)$$

（8·10）式中：P为难度（答对率）。

$\bar{X}$为全体考生某题实得分数的平均值。

a为该题的满分值。

（二）选择题难度公式之一

$$P=\frac{R}{N} \qquad (8\cdot11)$$

（8·11）式中：P为难度（通过率）。

R为选中某题正确答案的人数。

N为全体考生人数。

（三）选择题难度公式之二

$$P=\frac{P_H-P_L}{2} \qquad (8\cdot12)$$

（8·12）式中：P为难度（通过率）。

P_H为高分组选中正确答案人数的比例。

P_L为低分组选中正确答案人数的比例。

采用公式(8·12)计算难度，首先要排出考生在各题得分的序次，然后从高分一端依次取27%的考生为高分组，再从低分一端依次取27%的考生为低分组，最后用高分组选中正确答案人数的比例减去低分组选中正确答案人数的比例除以2，所得商数即是该题的难度指数。

以上三种公式求出的难度，实为易度，即难度指数越大的试题反而越容易. 为使难度指数与试题的实际难度成正比，名实相符，可用未通过率难度公式计算：

（四）未通过率概念之难度公式

$$f=1-\frac{R}{N} \qquad (8\cdot13)$$

（8·13）式中：f为难度（表示未通过率）。

R为选中某题正确答案的人数。

N为全体考生数。

（五）标准难度公式

由于用前文四种公式所求得的难度值，是答对、通过或未通过某题人数的百分比，仅表明一次考试每道试题难易度的顺序，而各题的难度不能相互比较。因而，最好采用标准难度公式计算试题的P值。标准难度值，是一种正态曲线下等距的难度值，优点在于各题难度可互比。其计算公式为：

$$\Delta=13+4x \qquad (8\cdot14)$$

（8·14）式中：Δ为标准难度。

4为标准差数。

x为标准差值。

13为试题标准难度的平均值。

试题的P值转换为标准难度后，其难度指数在1～25这一区间。若某题的P值为0.5，属于中等难度，按公式（8·14）转换成标准难度则为：$\Delta=13+4x=13+4(0)=13$，也正好是标准难度的平均值。若某试题的P值达到最高指数1，其标准难度也达到最高指数25，即$\Delta=13+4x=13+4(+3)=25$，说明该题无人通过或答对。如果某题的P值为0，则表明全体考生都通过或答对了此题，其标准难度便处于最低指数1的位置，即$\Delta=13+4x=13+4(-3)=1$。当试题的P值求出后，转换标准难度的程序比较简便，只要在正态曲线面积表上查出了P值的标准分数，就可直接代入公式（8·14）计算Δ值。

二、区分度分析

试题的区分度，是衡量试题鉴别考生水平差异能力的重要指标。进行区分度分析的目的，在于检验各道试题鉴别能力（和区分能力）的强弱，或试题是否能准确区分不同水平考生的差异，即实际水平高的考生得分高，实际水平低的考生得分低。

区分度用D表示，其指数在－1～1之间，即－1≤D≤1。D＝－1，说明考生的成绩与他们的实际水平刚好相反，水平低的考生得分高，水平高的考生得分低。D＝0，表明试题毫无区分能力，要么全体考生都得高分，要么所有考生全是低分，故无法判别应试者个体水平的差异。D＝1则表明试题的鉴别能力强，能够把不同水平的考生成绩准确区分开，水平越高的考生成绩越优。但这并不是说，区分度指数愈大的试题，其鉴别能力就愈强。实际情况是，鉴别能力最好的试题，其区分度指数一般在0.～05.65之间，区分能力较好的试题，其区分度指数多在0.4～0.5之间，试题区分度指数愈接近0鉴别能力愈差。

试题区分能力的检验，大都采用求试卷总分与各题得分相关系数的方法进行，其计算公式有如下几种：

（一）分组计算法

分组计算法的程序是，在排出考生各题得分次序后，从高分、低分两端各取27%的考生为高低两个组，并分别算出两组考生的通过率，再用公式（8·15）计算D值。

$$D = P_H - P_L \qquad (8\cdot15)$$

（8·15）式中：D为区分度。

P_H为某题高分组的通过率。

P_L为某题低分组的通过率。

（二）积差相关计算法

积差相关计算法，是通过求各题考生得分与试卷总分之间的相关系数，并以相关系数数值的大小表示试题区分能力的强弱。应指出的是，只有在考试信度较高的情况下，才适宜采用积差相关法求D值。在计算出考生的原始分数后，若要求试题的D值，可直接将分数代入公式（8·16）计算。

$$D=r=\frac{\Sigma XY-\frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\Sigma X^2-\frac{(\Sigma X)^2}{N}}\sqrt{\Sigma Y^2-\frac{(\Sigma Y)^2}{N}}} \tag{8·16}$$

（8·16）式中：D为区分度。

r为相关系数。

X为某考生某题实得分数。

Y为某考生试卷总分数。

N为考生数。

（三）点二列相关计算法

点二列相关系数用r_{pb}代表，按此法计算试题D值的公式为：

$$D=r_{pb}=\frac{\overline{X}_p-\overline{X}_q}{S_x}\cdot\sqrt{pq} \tag{8·17}$$

（8·17）式中：D为区分度。

r_{pb}为点二列相关系数。

$\overline{X}_p$为P部分X数列的平均值。

$\overline{X}_q$为q部分X数列的平均值。

p为二分变量中一个变量的比例。

q为二分变量中另一个变量的比例（q＝1－p）。

S_x为连续变量X变量的标准差。

点二列相关主要用于检验两列变量中二分变量对另一连续变量产生影响的一致性。若用点二列相关法计算试题的区分度，即是求考生各试题得分（通过或未通过、答对或答错的比例的量化指标）与试卷总分之间的相关系数，所得结果就是各题的D值。当两列变量中有一列是二分变量，如成功与失败、及格与不及格、答对与答错、通过与未通过等，而另一列变量属于连续变量时，就可用点二列相关计算法求试题的D值。同时还可用双列相关法计算试题的区分度，其计算公式为：

$$D=r_b=\frac{\overline{Y}_p-\overline{Y}_q}{S_x}\cdot\frac{p\cdot q}{r}.$$

如果欲求语文科考试某题的区分度，就以该科考生的分数作连续变量，以某题高分、低分两组的通过率或答对率为二分变量，再用双列相关计算公式求得。

在用区分度指标检验试题优劣时，应注意几点：一是在整个考试信度、效度极低的情况下，求取试题的区分度无实际意义，因为求区分度是计算考生各题实得分数与试卷总分的相关系数，如果考试总分不能代表全体考生的实际水平，区分度的量化计算也就在整体上失去基础。二是考试信度、效度高不等于试卷每一道题的区分度都好，进而对试题全部肯定。相反，考试的信度、效度低，也不能简单否定所有试题，而把少数鉴别能力强的题也给抛弃了。三是可以通过试题的P值估价试题的D值，一般说来，P值指数过大或过小，试题就不会有理想的D值，而P值适中的试题，往往鉴别能力较强。

三、信度分析

信度是检验考试质量的主要指标，也是各种考试必须注重的

目标，特别是选拔、合格、成就和研究性考试，更要追求高信度，否则，考试就不能如实反映考生的水平，使应试者得到公正的评价、准确的选拔和合理的任用，也就不能客观评估教学质量、给教育行政部门及劳动人事部门提供可靠信息，进而产生良好的社会效应。这样的考试，诚然不具科学性、公平性和教育性。但也必须明确，凡测量必有误差。所以，人类各种测量结果的可靠程度，只能用误差比重的大小进行衡量，而不能以有无误差为检验标准。正是在这个意义上，人们才将信度定义为真实分数方差与实得分数方差之比，亦即实得分数方差中误方差所占的比率。在考试分析方面，通常所讲的信度，是指用同一份试卷两次或多次测试条件基本相同的对象，所测结果相符合的程度。倘若各次测试结果大致相同，表明考试结果的信度高，测试结果一次一个样，就说明考试结果的可靠性差。在现行考试中，考试卷信度值一般要求达到0.9以上，因为信度值的最高指数为1，最低指数为0。信度指数达到1，意味着考试所测得的质和量没有丝毫误差，这在客观上不可能，故考试信度值的最高标准也只能定在0.9＜1的范围。至于考试的信度等于0，则表示考试测得的质与量完全是随机的。此种现象同样为考试实践中所罕见。

分析考试的信度，目前普遍采用的方法，是求两次考试结果或一次考试两半试题得分的相关。具体有以下几种：

（一）分半法

分半法是一种计算考试半分信度系数的方法，为现行考试广为应用。其公式为：

$$r=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}} \quad (8\cdot18)$$

（8·18）式中：r为两半试题成绩的相关系数(亦即信度系数）。

i为某考生。

X_i为某考生一半试题的实得分数。

$\overline{X}$为全体考生一半试题所得分数的平均值。

Y_i为某考生另一半试题的实得分数。

$\overline{Y}$为全体考生另一半试题所得分数的平均值。

n为考生数。

由于分半法是把一次考试的试题按奇偶题号和难度序次为等量的两半，并视两半试题的结果为用同一试卷对同等水平考生进行两次考试的结果，来计算一次考试的信度，所求信度系数只是半分信度系数，或谓同一试卷两部分试题得分的一致性系数。这种半分信度值一般低于全份试卷的信度值，因为考试信度与试题长度的关系甚为密切，题量大的考试信度通常要比题量小的考试信度高。所以，在半分信度值大于0小于1的情况下，应用公式（8·19）进行校正，并以求得的数值作为整个考试估计信度。

$$r_{xx}=\frac{2r}{1+r} \qquad (8\cdot19)$$

（8·19）式中：r_{xx}为整个考试的信度值。

r为半分信度值。

（二）等值法

采用等值法求考试的信度，同样用公式（8·18）计算。所不同之处在于，等值法是计算两份等值试卷得分的相关系数，而不是求两半试题得分的相关系数。等值法求出的两份试卷的相

关系数叫等值复本信度。它是用等值的两份试卷在恰当时间内对同一考试对象进行两次测试，再检验两次考试结果相符合的程度，两次测试成绩越接近，考试的可靠性越大。等值法虽可克服分半法的不足，但对试卷编制技巧的要求甚高，同时有两次考试间隔时间长短不易掌握、施测条件难于一致的缺陷。也就是说，要使两份试卷的标准和两次考试施测的具体条件完全相同，在考试实践中很难达到。

（三）再测法

再测法是用同一份试卷在不同时间对同一组考生施测两次，再用公式（8·18）计算两次考试分数的相关系数，然后以相关系数值的大小为据，判定考试的可靠性程度。或者说，以两次考试分数的相关系数值的大小作为考试信度的高低。

再测法、分半法和等值法，虽然都是用计算两组分数相关系数的方法来检验一次考试信度的高低，再测信度、估计信度、等值复本信度的意义并无本质区别，计算公式也相同，但三种方法的不同点是极为明显的：其一，再测法是以同一份试卷两次测试相同对象的分数作为求相关的两列变量，等值法则以两份试卷在不同时间测试同一对象的分数作为求相关的两列变量，而分半法是把一份试卷两半试题的得分作为求相关的两列变量。其二，再测法和等值法用公式（8·18）求出的相关系数，可直接作为一次考试的信度，分半法用公式（8·18）求得的相关系数大于0小于1时，还需用公式（8·19）校正后，才能对一次考试的信度作出估价。其三，尽管三种方法都可检验考试的信度，但检验的角度各不相同。再测法是从考试结果稳定性的角度进行检验，而分半法和等值法则是从两半（或两份等值试卷）试题测试结果的一致性角度进行检验。其四，三种信度计算方法的适应范围不同，等值法适合于评价应试者在较大范围内所具知识、能力之考试的信度计算，而以预测应试者学习或工作能力为目的的考

试，如招工，招干和招生考试等，则采用分半法或再测法求考试信度更为适宜。此外，若考试使用的是二重选择题，最好用库德—理查逊公式 $r_x = \frac{n}{n-1}(1 - \frac{\sum pq}{S_x^2})$）计算信度值。而多重选择题和问答题考试的信度值，又宜用克郎巴赫 α 系数公式（$\alpha = \frac{n}{n-1}(1 - \frac{\sum S_1^2}{S_x^2})$））计算。

四、效度分析

考试效度是指考试的有效性程度，也就是考试已测到的质和量与主试者欲测的质和量相符合的程度。一般用考试结果（考试分数）与能体现考试目的的效标分数之间的相关系数（r）表示。因此，分析考试的效度必须首先确定效标。所谓效标，即是根据不同考试目的确定的能够代表考试目的的效度标准，反映这种人为确定的效度标准的分数，称为效标分数。考试分数与效标分数两列变量之间的相关程度，大致表现为五种情形，如图8－4所示。

图8－4表明相关系数的值域为 $-1 \leqslant r \leqslant 1$。$r=1$，表示完全正相关，说明测试结果完全符合欲测的目的。$r=-1$，表示完全负相关，说明测到的内容与考试欲测的内容彻底相反。$r=0$，表示0相关，说明考试结果与欲测目的无关。$r \approx +0.5$，表示正相关，说明考试结果与欲测目的趋向一致。$r \approx -0.5$，表示负相关，说明测到的内容与欲测的内容呈反向趋势。

在现代考试或测验中，效度作为检验考试质量的主要指标使用极广，且因各种考试目的有别，加上每一种效度又有特定的检验范围，所以效度名目繁多，诸如内容效度、预测效度、协同效度、结构效度、共时效度，假设效度，等等。美国心理学界曾把各种考试的效度归并为三大类：内容效度，效标关联效度和结构

效度。下面概要介绍考试质量分析使用较多的两种效度——内容效度和预测效度。

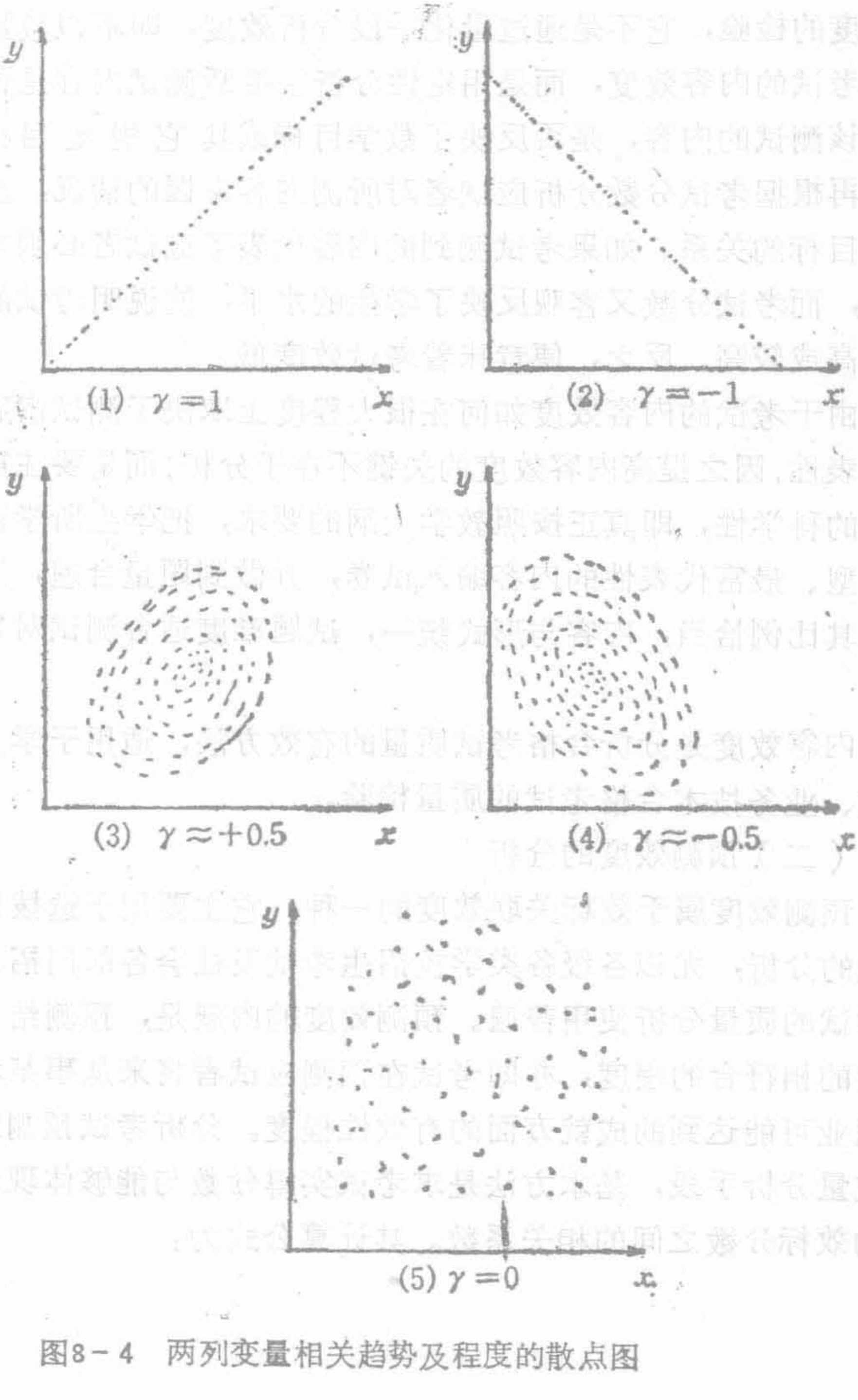


图8－4 两列变量相关趋势及程度的散点图

（一）内容效度的分析

分析考试的内容效度，意在检验考试已测到的内容是否符合考试需要测试的内容。检验考试内容效度，其方法有别于其它考试效度的检验，它不是通过量化手段分析效度，即不以数量指标估价考试的内容效度，而是用定性分析法推断测试内容是否代表了应该测试的内容，是否反映了数学目标或其它规定目标的要求，再根据考试分数分析应试者对所测内容掌握的情况，及其与既定目标的关系。如果考试测到的内容代表了应试者必须掌握的内容，而考试分数又客观反映了学生的水平，就说明考试的内容效度高或较高。反之，便意味着考试效度低。

由于考试的内容效度如何在很大程度上取决于测试内容本身的代表性，因之提高内容效度的关键不在于分析，而是要注重试题设计的科学性，即真正按照教学大纲的要求，把学生所学范围内最典型、最富代表性的内容编入试卷，并做到题量合理，所用题型及其比例恰当，内容与形式统一，试题难度适合测试对象的实际。

内容效度是分析合格考试质量的有效方法，适用于学业成绩考试、业务技术合格考试的质量检验。

（二）预测效度的分析

预测效度属于效标关联效度的一种，它主要用于选拔性考试质量的分析，尤以各级各类学校招生考试及社会各部门招工、招干考试的质量分析使用普遍。预测效度的内涵是，预测结果与考试目的相符合的程度，亦即考试在预测应试者将来从事某种学业或职业可能达到的成就方面的有效性程度。分析考试预测效度采用定量分析手段，基本方法是求考试实得分数与能够体现考试目的的效标分数之间的相关系数。其计算公式为：

$$r_{xy}=\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{N\cdot SD_x\cdot SD_y} \qquad (8\cdot20)$$

在（8·20）式中：r_{xy}为预测效度。

X_i为某考生的考试分数。

$\bar{X}$为考试分数的平均值。

Y_i为某考生的效标分数。

$\bar{Y}$为效标分数的平均值。

i为某考生。

SD_x为考试分数的标准差。

SD_y为效标分数的标准差。

用计算考试分数与效标分数相关系数的办法检验考试的预测效度，不能以求得的相关系数值的大小直接决定预测效度的高低。当用公式（8·20）求出相关系数后，还应查对相关系数表，看所得数值是大于、等于还是小于表中的临界值。然后根据查对情况确定考试的预测效度。若求得的相关系数值大于或等于表中的临界值，表示考试分数与效标分数有显著相关或高相关，说明考试的预测效度高。若小于临界值，说明两种分数之间无明显相关，不相关或完全相反，便可得出考试预测效度低或预测结果无效的结论。

在分析考试的效度时，有三点是必须明确的：一是不能以考生分数的高低作为判定考试有效性程度的标准。考生在某种考试中取得的分数，往往是多种知识综合反映的结果，不单是所测方面知识或能力水平的反映。即使一次考试考生全得高分，但主试欲测的内容很少测到，甚至测到的内容与欲测的内容相反，考试的效度仍然是低的，因为考试结果很少或未能实现预期的目的。二是不能以考试信度指数的高低为标准确定考试的效度。因为信度的

高低只表明测到的内容是否反映了考生的实际，而并不反映考试所测得的内容是否是主试者欲测的内容。检验考试效度，只能以测得的内容与欲测的内容之间的一致性程度为标准，测得的内容愈接近欲测的内容，考试的效度愈高。三是考试的有效性是针对一定的考试目的而言的，一次考试只对某一考试的具体目的有效，而不是对各种考试目的都有效，比如汽车司机驾驶执照的考试，目的是检测被试有无驾驶汽车的知识与实技，其考试的有效性就只能针对这一设考目的进行检验，而不能用是否测出了被试驾驶飞机的知识与实技为标准，来检验该次考试的效度。

第九章

考试的组织机构及其职能

健全而又富有活力的考试组织机构，是国家推行考试制度的组织保证，也是对整个考试流程实现科学控制的必备条件。

在漫长的考试实践中，人们对考试所苦心研讨的无非是三个问题：探讨考试本身的规律，提高制作考试这把量尺的精确度；摸索使用考试这一量具的科学方法，提高考试结果的客观性；研究考试的固有功能，增强考试主动适应社会发展需要的能力。对这些问题认识的深度，以及科学解决这些问题所达到的程度，总是与一个国家考试组织机构的状况处于同步关系。这是因为：没有健全的考试组织机构，有关考试方针、政策、法规的制订，就很难建立在深入调查、细致研究、严密论证的科学基础之上；没有健全的考试组织机构，即使有科学的考试制度，也难以得到圆满的贯彻实施；没有健全的考试组织机构，就不可能有稳定的专业考试队伍，整个考试的设计与实施必然是临时凑兵，考试量尺标准、实施规范、结果准确的目标，实难期及；没有健全的考试组织机构，自然也谈不上深入持久的考试科学研究，并随社会对考试要求的提高不断更新、完善考试的理论，用以指导新的考试实践，进而强化考试主动适应社会的能力，使之正确发挥督导功能与调节功能。所以，当人类社会进入20世纪中期以后，由于

社会生产力的迅猛发展，以及新技术革命的浪潮对旧有社会观念的剧烈冲击，从而引起了“知识蓄积主义”人才观、“仓库理论”教育观和偏重知识记忆检测的考试观的深刻变革，以及考试自身的发展。一方面，社会要求考试迅速摆脱传统观念的束缚，尽快形成适应社会发展需要的考试模式，考试的设计与实施，必须符合科学性、客观性、公平性、实践性的原则，做到量才客观、选才准确、信息可靠。另方面，考试观念的更新及考试门类的扩大，又导致考试理论、考试程序、考试手段的革命，使考试逐步成为一门科学，考试流程成为高度专业化的科学控制过程。与之相适应，一些国家和地区的考试组织机构日益整体化、专业化，形成了完整的专业考试系统。而且考试机构及其队伍的规模年有扩大，并处在不断发展的态势。正是从考试科学化程度与考试组织机构完备状况同步发展关系的角度，人们才把考试机构是否系统化、专业化，作为一个国家教育、考试、人事管理是否现代化的重要标志之一，也正是由于考试组织机构在现代国家管理中的重要作用，人们才注重考试组织机构的建立与健全。

第一节　考试机构与人员结构

考试组织机构，它作为国家测度、甄选各类后继人员、新从业人员和在业人员的职能部门，必须依据国家管理及事业发展的实际需要[illegible]。因此，不同社会制度和经济水平的国家，有不同的考试组织机构沿革，在同一历史发展阶段，各有不同的考试组织机构模式。至于人员结构，又取决于考试组织机构所负任务的需要，各国之间乃至各考试机构之间，也从无统一的模式。基于上因，对于考试机构与人员结构的讨论，只能通过典型分析而述其共性，有关现今世界五花八门的考试组织机构模式及其人员

结构情况，拟在《中外考试史》中具体探讨，此不一一列举。这是首先要说明的一点。

一、考试的组织机构

如前所述，一个国家的考试组织机构是因国家管理的实际需要而设立，并随社会发展对考试需求的变化，以及考试自身的发展而演变的。从现阶段不同国家和地区的考试组织机构系统来看，可从考试组织机构的隶属关系、行使职能的期限、职能性质三个角度，分为七种类型。

（一）从隶属关系的角度划分

按照考试组织机构的隶属关系，现阶段不同国家的考试机构可分为两大类：

1．独立经营的考试机构。

这种考试机构产生于本世纪40年代，它与政府统属的考试机构并存，是一种独立的专业考试机构，专司考试设计、实施和研究，属业务性质。其基本特点是：由民众团体或学校自发集资兴办，无需政府拨款；单独行使职能，不受政府干预；在考试实施过程中，出现问题由考试机构自行处理，与政府概无关联；以所主持考试的质量维持考试机构的存在和发展。目前，这类考试机构有国际性的、国家性的，也有地区性的，承担的考试业务多种多样，既设计和主持定期举行的大规模考试，如国际、国家或某一区域性的考试，也可设计和主持不定期的、规模较小的考试，象学业成绩、职业技能、某些研究性考试等。概要地说，这种独立经营的考试机构，是以设计、主持、评判、研究考试为职业，以给考试结果使用单位提供客观可靠的依据为目的。

独立经营考试机构为其性质和职能所决定，机构内部一般结构完整、部门齐全、分工细致。如当今世界上规模最大的独立考试机构——美国教育考试服务公司，就设有各种试题编审委员

会、秘书处、科研所、印刷所、统计分析中心、现场服务处、行动计划处、法律顾问处、对外联络处、财务审计处、技术情报中心等职能部门。该公司创建于1947年，目前拥有职业工作人员2700余人，在世界156个国家设有考点近10,000个，每年参加他设计和主持的各种考试的考生约7,000,000人。其组织机构系统如图9-1所示。



图 9-1　美国教育考试服务公司组织机构简图

又如地区性的独立考试机构——香港考试局。它建立于1977年，现设有会考委员会、会考执行委员会、17个科目协调委员会、财政及人事委员会、秘书处及其下属的6个组：发展组、总务组、会考组、系统组、财务组和海外考试组，详见图9-2。拥有职员160余人。该局除负责主持香港中学会考、香港高等程度会考和香港高级程度会考外，还承担海外（主要是英国和美国）各种类型和规模的考试。

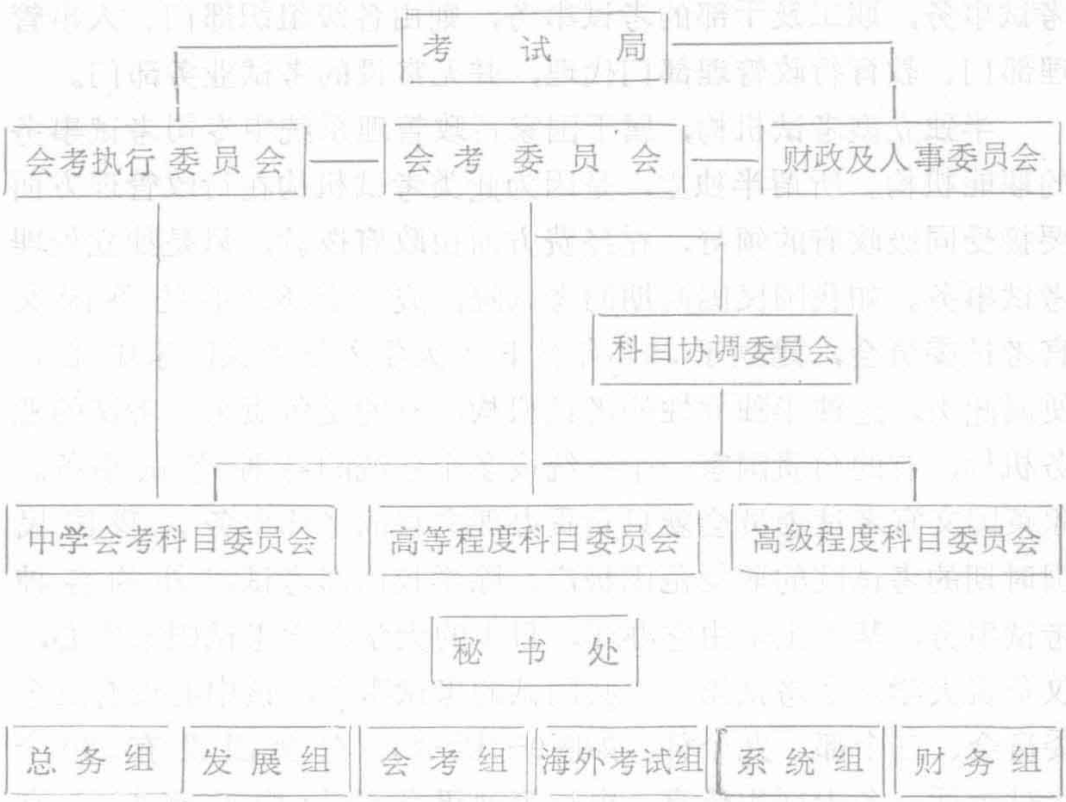


图 9-2　香港考试局组织机构简图

2．政府统属的考试机构。

政府统属的考试机构，历史上早已有之，并非现代社会的产物。此种考试机构总体上属于国家行政管理系统的职能机构，负责国家考选事务的决策和实施。若按考试机构设置的形式，又可分为附属型考试机构和半独立型考试机构。

附属型考试机构，一般分设在国家有关行政部门，或由国家其他职能部门代理考试业务，没有完整的组织机构系统。比如我国封建时期的考试事务，就由礼部、吏部等代理。美国文官考试任用业务，由美国文官委员会办理。现阶段我国教育系统的考试机构，大都分设在各级教育行政部门，如普通高等学校招生办公室、自学考试办公室、成人考试办公室、中等学校考试办公室等，均属各级教育行政部门的分设机构，而各级学校教育的内部

考试事务，职工及干部的考试事务，则由各级组织部门、人事管理部门、教育行政管理部门代理，并无常设的考试业务部门。

半独立型考试机构，属于国家行政管理系统中专司考试事务的职能机构。所谓半独立，是因为此类考试机构在行政管理方面要接受同级政府的领导，在经费方面由政府拨款，只是独立处理考试事务。如我国民国时期的考试院，成立于1855年的英国文官考试委员会，建立于1977年的日本大学入学考试国家中心，便属此类。这种半独立性的考试机构，有的是负责某种考试的业务机构，有的负责国家一个系统或多个系统的各种考试事务。象英国文官考试委员会就只负责办理文官的考任事务。我国民国时期的考试院的职能范围极广，除学校内部考试以外的各种考试事务，基本上全由它办理。日本的大学入学考试国家中心，仅负责大学入学考试第一次共同试的考试事务，该中心设有三个委员会、三个部、九个科，如图9-3所示，在全国设有280个考试大厅，各大厅用传真、电话等通讯系统与中心连接，有5640个考试专用房间，每年参加大学入学第一次共同试的考生约35万左右。日本大学入学考试国家中心现有固定职员近100人。

（二）从职能性质的角度划分

从职能性质的角度分，考试机构可分为三种类型。

1．决策性考试机构。

决策性考试机构，以制定考试方针、政策、法规，审定各种考试标准、方案，核准各类考试执行机构，督导和检查考试制度的贯彻执行为主要职能，少数还兼有考试行政管理的职能。这种考试机构通常设在国家人事机关、教育行政管理机关和组织部门，很少单独建制。如英国由文官事务部代理，美国的“政策和标准局”、“录用和考试局”设在文官委员会，如图9-4所示。日本由内阁人事院代理。在我国则由国家教委、组织部、劳动人

事部和其它部委及下属的各级政府职能部门的有关司（局）厅、处、科代理。

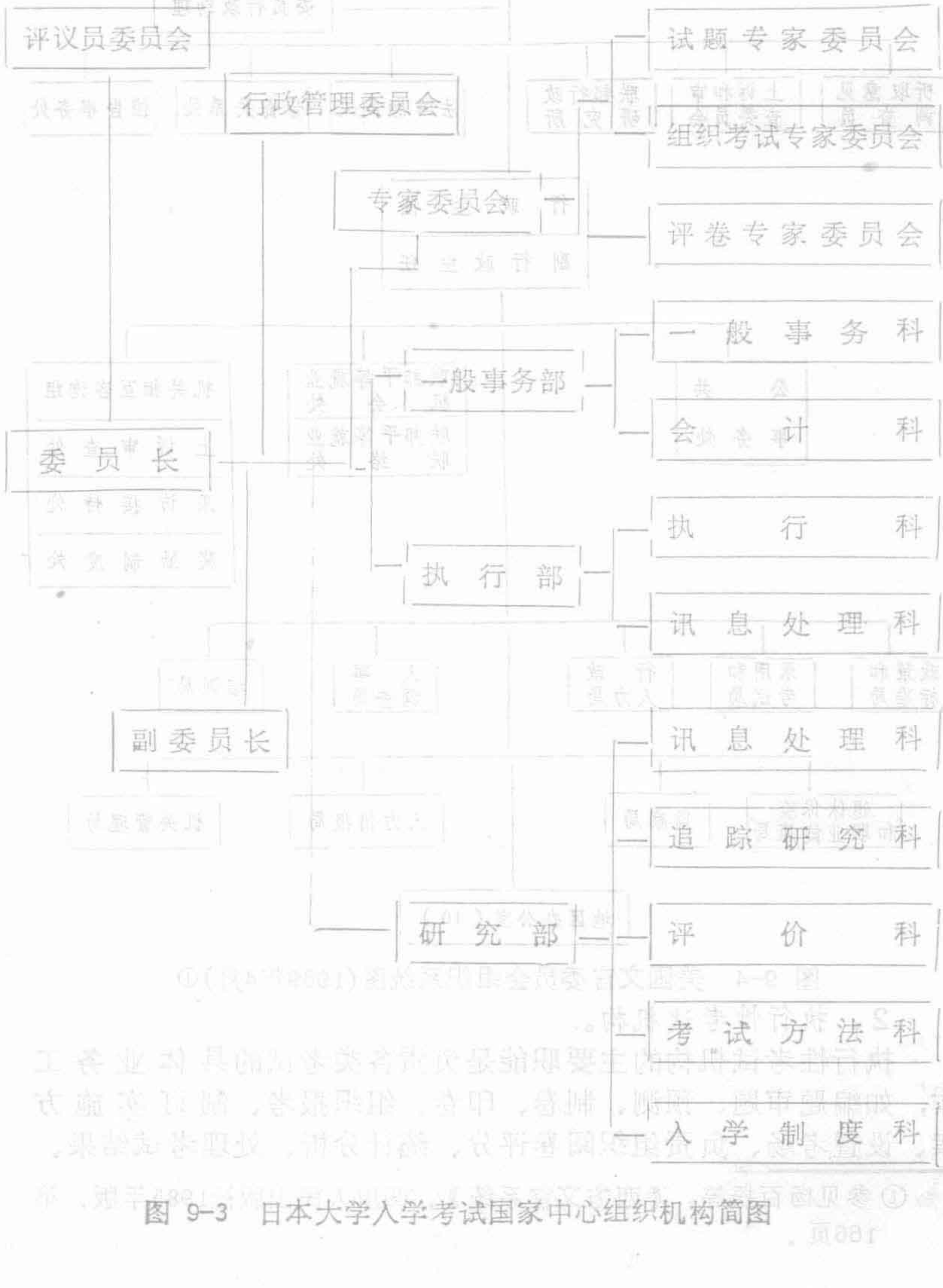


图 9-3　日本大学入学考试国家中心组织机构简图

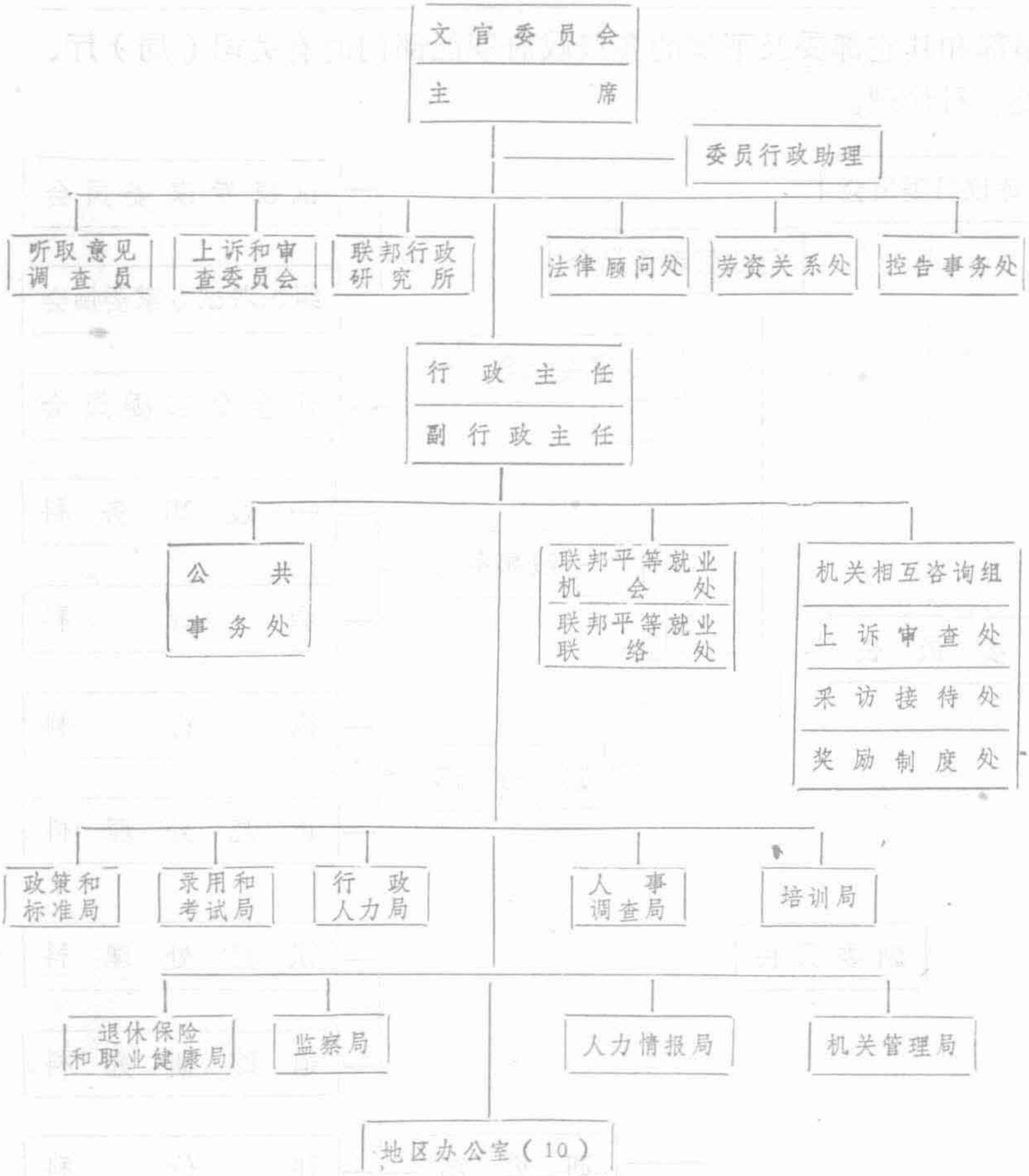


图 9-4 美国文官委员会组织系统图(1969年4月)①

2．执行性考试机构。

执行性考试机构的主要职能是负责各类考试的具体业务工作，如编题审题、预测、制卷、印卷、组织报考、制订实施方案、设置考场、负责组织阅卷评分、统计分析、处理考试结果、

① 参见杨百揆等：《西方文官系统》，四川人民出版社1985年版，第166页。

开展考试研究，等等。有些执行性考试机构，还负责颁发考试合格证书、培训考试专业人员。象日本的大学入学考试国家中心，我国及其他国家在教育、人事、组织等政府行政管理系统中设立的业务考试机构，均属执行性考试机构的范畴。

3．综合性考试机构。

此种考试机构，属国家单独建制的考试行政事务机关，它负有决策性和执行性考试机关的双重职能，既是国家考试系统的权力机关，又是承办各种考试事务的业务机构。美国的文官委员会，我国民国时期的考试院等，便属综合性的考试机构。我国现阶段的各级招生办公室、考试办公室，也在一定程度上兼有决策与执行的综合性质，它不是纯粹执行性考试机构，又不负有决策性考试机构的全部职能。

（三）从行使职能的期限划分

按照行使职能的期限，考试机构可分为两大类。

1．常设考试机构。

这类考试机构属国家法定的考试行政或业务部门，长期行使规定范围内的职能，除非常时期外，一般稳定不变。上述五种类型的考试机构，属于常设考试机构的范畴。无论是政府统属的考试机构，还是独立经营的专业考试机构，均需得到国家有关部门的核准后，才能行使其职能。两者的区别在于，政府统属的常设考试机构，是代表政府行使职能，民众团体兴办的独立考试机构纯属民间性质，是一种不反映政府意志、经济独立、只承办具体考试业务的常设考试机构。

2．临时考试机构。

临时考试机构一般因工作急需而设立，随工作的结束而撤消，机构、人员、任务都是临时性的，其职能单一，理职期限短。如为测度、甄选某种特殊人才、评定学术职称学位等，而设立的论文答辩委员会和职称考评委员会，就是一种临时性的考试

机构。

二、考试机构的人员结构与素质

考试机构的人员结构问题，实际上是考试机构内部单位结构（下属职能部门）与职位结构问题。一般来说，一个考试机构有什么样的单位和职位结构，就有与之相适应的人员结构。当然，就总体而言，考试机构的人员结构必与其机构的目标要求相一致。因而，不同目标的考试机构有不同的单位结构和职位结构，进而也就有不同的人员结构，并不是凡考试机构都有相同的人员结构。下面着重探讨专业考试机构的人员结构问题。

（一）确立专业考试机构人员结构的基本原则

1．紧扣机构服务目标。

人员结构应与考试目标保持一致，这是专业考试机构人员设置的基本原则。比如美国教育考试服务公司、日本大学入学考试国家中心、香港考试局三个专业考试机构，由于都是以设计、主持教育系统的考试为机构的主要目标，故在人员结构上有许多共同点，但因各自的具体目标不尽相同，其机构内部的单位结构和职位结构又有明显的区别；美国教育考试服务公司设计、主持的考试项目，纵含多种层次，横涉众多学科，它除了主持从小学至研究生院各个层次的考试外，还主持各类教师的资格审查考试，以及50多种职业的命题和资格审查考试。因此，它的人员结构就呈现出两多的特点，仅以2700余名职员中的专家或高级专门人才结构为例：一是人员层次多。近700名高级专门人才，分为三大层次：第一层次是175名获博士学位的高级专家；第二层次是245名获硕士学位的各类学科专家或技术专家；第三层次是200余名具有学术、技术专长的高级专门人才。二是人员类别多。在近700名专家和高级专门人才中，除教育学、心理学、教育测量学、心理测量学、统计学、数学、计算机科学方面的专家外，还

包括若干门类的学科专家、技术专家和管理专家。至于一般管理人员和技术人员，同样涉及多个方面的多个层次。日本的大学入学考试国家中心，因其服务目标单一，仅主持大学入学第一次共同试，所以，它的人员结构远不象美国教育考试服务公司那样复杂。如中心内部高层次结构的三个委员会，即评议委员会、行政管理委员会和专家委员会，均由国立大学校长、教授、教师、学识或经验丰富的人士及中心的教授组成，不涉及其它系统或层次的人员。大学入学考试国家中心内部的固定人员结构，也较美国教育考试服务公司简单，如表9-1所示。香港考试局的人

表9-1　日本大学入学考试国家中心人员结构表

所长	副所长（计划调整官）	管理部、事务部	研究部			合计
		技官、事务官等	教授	助教授	助手	
1	1	73	5	5	4	89

员结构又有别于上述两个考试机构，它是以中学会考、高等程度会考和高级程度会考为主要服务目标，其人员结构主要包括大学、理工学院、专门学院、中学的高层人员及公职人士与工商界知名人士等，而且中学教师占有较大比例。如在中学会考、高等程度会考、高级程度会考三个科目委员会的54名成员中，中学教师占21名，为三个委员会成员总数的38.9%，为全局人员总数的13%。

2．强化机构主体。

专业考试机构的主体，系指那些影响机构目标实现、关系机构生存与发展的关键性职能部门。从专业考试机构的基本职能看，影响机构目标实现的职能部门，最主要的是考试设计部门和考试实施部门。再者，考试质量是决定考试机构生存与发展的根

本存在，而要确保机构所主持考试的质量，就不能没有统计分析和考试理论、方法、技术手段方面的职业研究部门。这就是人们通常所说的决定专业考试机构生命的三支队伍。也就是说，这三个职能部门，在任何一个专业考试机构中都是主体，必须置其于单位设计和人员配备的首要地位。机构其它辅助性部门的设置及人员配备，也要以有利于这三个主体部门的密切合作、协调运转、产生最佳群体效能为原则。

（三）人员的结构与素质

根据专业考试机构的基本职能，以及专业考试机构单位结构和人员结构设计的一般原则，专业考试机构的整体人员结构，当由考试设计、考试实施、考试研究与质量分析、后勤服务四方面的人才组成，并要职位结构合理、各类人员比例恰当，每个职员都有履行所负职责的素质和能力，使机构的部门与部门、人员与人员之间能够相互感应和互补，有利于群体功能的发挥。对各类人员具体结构与素质的要求是：

1．考试设计人员的结构与素质。

考试设计部门是专门考试机构的三大主体部门之一，亦即目前国外一些专业考试机构中常设的试题编审委员会（或试题专家委员会、科目委员会）其主要职能是制作考试的标准量尺，即编题、审题、选题、拼配标准试卷和等值复本。如果说考试量尺标准是考试成功的第一要素，那么，考试设计人员的质量优劣就是决定考试机构能否实现既定目标的首要条件。因此，对考试设计人员的配备必须慎重选择，他们必须是具有良好理论素养、较深专业造诣和受过严格业务训练的专家或专门人才。具体要求是：

在结构上既要精又要全。所谓精，即是负责考试设计的人员要精干，不宜数量太大，应人与事相配、能与职相称，切忌因人设庙。所谓全，包括两层意思：一是人员的专业结构要全，即凡

考试机构所主持考试的每一门类（或涉及的学科、专业）都要有相应的设计人员负责本类或本门学科的试题编审，不可跨门类或跨学科设计；二是人员知识、智能结构要全，考试设计是一项综合性很强的工作，单靠一方面的人才难以成就，一份科学的试卷总是多方智慧交互作用的产物，所以，考试设计人员应有较为完全的知识、智能结构，至少要有四方面人才的密切合作，即考试、教育、心理和各门学科的专门家。就层次、比例讲，考试、教育、心理方面的专家居决策指导层，比例约占设计人员的1/4，各门学科专家居实施层，比例为设计人员总数的3/4。

在素质方面，凡参与考试设计的人员必须具备三个懂得、一个精通。首先，要懂得考试学、教育学、心理学、统计学、人才学及行为科学的基本理论与知识。其二，懂得不同考试对象的生理、心理特征，不同类型考试的性质与特点，不同类型试题的功能。其三，懂得考试的基本规律与原则，熟知考试设计的程序与要求，有较高的命题实技。其四，透彻了解本门学科的教学大纲和教材内容，精通本门学科的专业知识。同时，要有服从真理、善于合作的品质，在考试设计过程中，能虚心听取不同意见、勇于正视试题编审方面的缺陷，[illegible]作协调一致，能严格遵循命题的原则和程序，科学地编题、审题、选题和制卷。所编试卷应经得起各种质量指标的检验，标准稳定客观，能够维护考试机构主持之考试的历史继承性和均衡性。

2．考试实施人员的结构与素质。

要实现考试机构的目标，仅有精干的设计人员及其所制作的标准量尺，这只是提供了成功的可能，要使标准量尺产生良好的效应，并获得预期的结果，还需有能够正确使用标准量尺的考试人员付诸实施。正由于此，当今国外一些专业考试机构才把考试实施部门视为重要职能部门，如美国教育考试服务公司的行动计划处和现场服务处，日本大学入学考试国家中心的组织考试专家

委员会和执行部，无不居于考试机构主体地位，所需人员都经过精心选择。

考试实施人员的基本职能是，制订考试规程和实施细则，负责考试实施过程中各项工作的组织与管理。这部分人员为数众多，在机构人员总体中占有相当大的比例，且需雇请大批编外人员担负监考、阅卷等工作。为考试管理与实施的任务所决定，考试实施人员应由专职与兼职两部分人组成。其中专职人员包括两个层次：组织考试实施的测试专家和评卷专家；富有组织管理能力、考试业务娴熟的技术人才和管理人才。兼职人员主要是监考和阅卷的工作人员，他们在专职人员的指导下，完成职责范围内的任务。兼职人员除少数作为固定雇请对象外，大部分可根据工作需要临时选聘。

在素质方面，所有参与考试实施的人员，都必须经过一定的业务训练，具有高度的责任感，熟知考规考纪及实施程序，有准确执行考试规程、恰当处理考试过程中出现的各种事端的能力，以及根据不同考试对象的心理特征创设良好应试环境的能力，保证考试实施的各项工作能严格按照科学程序进行，能有效控制各类无关因素的干扰，做到实施规范、判分客观、结果准确。

3．考试统计、研究人员的结构与素质。

由考试数理统计、质量分析、科学研究和资料情报人员构成的考试研究与质量分析部门，是专业考试机构的三个主体职能部门之一。该部门的人员结构及效能如何，对考试机构的社会信誉影响极大，并在一定程度上决定考试机构的生命力，历来为专业考试机构所重视。如美国教育考试服务公司的科研所和统计中心，日本大学入学考试国家中心的研究部，香港考试局的系统与统计组，以及其他国家考试机构中的评价中心和统计研究处等，全是为考试数据统计分析、考试质量鉴定、考试理论、内容与方法改革探讨、考试资料情报信息收集、整理与研究等，而特设的

专门机构。

由于考试研究与质量分析部门的工作涉及面广，从微观讲，它要研究考试的内在结构，考试的程序与原则，考试的内容、方法和手段，题库理论与试题理论，各门学科的内容与特点，以及古今内外考试经验的比较，要负责各类考试数据的统计、整理和分析，各种信息的输入与输出。在宏观方面，需探讨考试的功能，考试与政治经济、文化教育、科学技术、家庭、考生诸方面的关系，分析影响考试质量的各种因素，并为考试机构提高考试的科学性与权威性、扩大服务目标，提供理论依据或改进方案；为评价、选拔、任用人才提供客观标准，为使用考试结果的部门输送可靠信息。如此复杂精细而又关系重大的工作，必须由一批训练有素的人员来承担。所以，考试研究与质量分析部门人员的结构成分较为复杂，既要有数理统计、数学、电脑、考试、教育、心理、情报方面的专家以及不同年龄、专业结构的研究人员，又要有程序设计、电子运算方面的专业技术人员，还要有资料情报人员等。

考试研究和质量分析人员的素质要求也比较高，他们不仅要具备本职工作所需的理论知识和实际技能、求实的工作态度和强烈的责任心，而且必须有探新精神和创造意识。

4．后勤服务人员的结构与素质。

这部分人的主要职能是为考试机构的正常运转创设物质条件。虽其所属部门在考试机构中处于辅助地位，对考试机构目标的影响不如三大主体部门直接和强烈，但不能因此看轻后勤服务部门在考试机构中的作用。事实上，后勤服务部门的工作与其他部门的工作效率息息相关，没有后勤人员的全心服务，就不可能有全体人员的饱满工作热情和旺盛的工作精力，工作效率必受影响。反之，即使其他部门人员的工作积极性很高，但无后勤部门提供必要的物质设备，创设良好的工作环境，也会事倍功半，欲

速则不达。因此，在专业考试机构中，后勤服务人员非但不可少，且应数量足、素质好。

以上四方面的人员，对于一个专业考试机构来说，可谓缺一不可。仅有考试设计人员而无考试实施人员，试题便不能生效；仅有考试实施人员而无考试设计人员，实施部门的职能又无法实现；仅有设计、实施人员而无统计分析人员，考试质量不能做出科学的结论，必然影响考试的社会效应；单有三大主体部门成员的努力，而无后勤部门的积极配合，也会因物质设备的短缺和工作条件的限制，而导致整个机构工作运转不灵。

最后还应强调指出的是，专业考试机构三个主体部门的成员，除应有良好的业务素质外，还必须有优良的政治素质，即要有崇高的考德，能严守考纪、秉公办事、不徇私情，坚持考试面前人人平等。

第二节 考试机构的职能

考试机构的职能取决于考试机构的性质与目标。比如决策型考试机构，就不负有执行型考试机构的职能，而综合型考试机构，又负有决策型与执行型两类考试机构的双重职能。因此，对考试机构的职能不可一概而论，应根据各类考试机构的性质及具体服务目标确定。但为了对考试机构的职能有一个比较全面的讨论，现将各类考试机构的职能归结为五大职能进行综述。

一、制定考试方针、政策及法规

科学的考试方针和政策、健全的考试法规，是一个国家维护考试严肃性和权威性、加强考试事务管理、保证考试制度顺利贯彻实施的必要条件。因此，自从有了国家，也就有了与其社会制

度相一致的考试方针、政策及法规。制定考试方针、政策及法规，便逐步成为决策型考试机构或综合型考试机构的基本职能，有的国家则由国家行政机关或立法机关代理。现阶段一些独立经营的考试机构，为做到依法施考，同样兼有制定各种考试规程的职能。属于制定考试方针、政策及法规这一职能范围的工作任务是：

（1）决定考试的组织机构及其职责；

（2）决定考试实施的形式、范围与原则；

（3）决定享受考试权利的条件；

（4）制定考试机构管理条例；

（5）制定各类考试的标准及甄选人才的原则；

（6）制定考务条例；

（7）制定考试档案管理条例；

（8）制定考试违纪处罚条例。

对以上内容，不仅应拟出文字章程，而且以政策或法律形式公诸于世，作为施考、应考及使用考试结果等各方人员必须遵守的准则。

二、负责考试设计

考试设计是专业考试机构或执行型考试机构的基本职能之一。考试设计的具体任务有：

（1）根据客观需要确立设考目的；

（2）制定设考项目的考试大纲；

（3）收集、整理、分析、研究考试内容的素材；

（4）编制试题设计蓝图；

（5）编审试题；

（6）组织预测；

（7）编制标准试卷、等值复本和标准答案；

（8）编写试卷使用指南；

（ 9 ）建立试题库。

为考试机构提供标准量具，是考试设计的本质任务。至于考试机构采取何种形式实现考试设计的职能，则无统一之规，可由考试设计部门独立完成，也可内外结合，分散进行。

三、组织考试实施

专业考试机构或执行型考试机构的另一基本职能，即是组织考试实施。由于专业考试机构所组织的考试一般规模较大、涉及面广，需动用大量人力，物力，其职能的实现，常常要借助外界力量。所以，组织考试实施的职能范围应包括：

（ 1 ）制订考试实施方案；

（ 2 ）组织报名；

（ 3 ）颁发准考证；

（ 4 ）组织监考、评卷队伍；

（ 5 ）负责试卷、答卷的印制、保管与发送；

（ 6 ）设置考场、试场；

（ 7 ）组织施测；

（ 8 ）负责试卷评阅；

（ 9 ）收集整理试场动态；

（10）负责处理违纪事件；

（11）编写考试分析报告；

（12）印制考试成绩手册；

（13）处理考试结果。

四、培训考试队伍

考试机构把培训考试队伍视为重要职能，主要出于两方面的需要：一是保证考试质量的需要，二是延续和发展考试机构生命力的需要。因为没有高质量人员组成的职能部门，以及由其形成

的合理整体结构，考试机构的效能就得不到正常发挥，考试质量必然降低。同时，考试机构人员的新陈代谢，既是自然现象，也是考试机构发展的必备条件，而对新补充的人员进行必要的职业训练，又是缩短新人员适应本职工作时距的有效措施。况且随着科学技术的发展，老从业人员也有一个再训练再提高的问题。培训考试队伍的主要职能为：

（1）培训考试机构在业人员；

（2）培训新进人员；

（3）培训监考人员；

（4）培训评卷人员；

（5）培训考务管理人员；

（6）培训统计分析及其他业务人员。

五、研究考试科学

人类长期的考试实践表明：没有科学的考试理论指导，就不可能有成功的考试实践；缺乏理论根基的考试机构，也就不可能有顽强的生命力。现今一些专业考试机构之所以把考试研究部门置于主体职能部门的地位，也正是前人成功与失败之经验教训启迪的结果。但这些专业考试机构也并不是一开始就把考试作为一门科学研究的，早先多限于考试内容、方法、技术问题的研究，真正研究考试的规律及系统的科学理论，是在新技术革命兴起以后，近年来有了长足的发展，研究的深度与广度远远超出方法、技术的范围，而是涉及考试现象的众多方面，诸如：

（1）考试产生、发展、演变的成因；

（2）考试与人类社会的关系；

（3）考试与人的身心特征发展；

（4）考试与教育的关联；

（5）考试的社会价值；

（6）考试模式的数理理论；

（7）试题反映理论及其应用；

（8）题库建设的理论与方法；

（9）考试的程序与原则；

（10）考试质量控制的理论与法则；

（11）考试结果量化的依据；

（12）考试数据的整理与分析。

第十章

考试与教育

考试与教育是人类两种独立的社会活动。它们因社会生产与社会生活的客观需要而产生，各以不同方式满足社会需求为归宿，始终是人类社会进化的重要机制。

综观人类社会发展的历史，没有全无教育活动的社会，全无考试活动的社会也同样没有。“从有教育以来就有考试的制度”。[1]只要有学校教育和社会分工的存在，考试就不会消亡。考试和教育始终是人类社会母系统中相互独立而又密切关联的两个子系统。

所谓独立，是因为两种活动性质有别、职能各异：教育是施教者与受教者两种主体之间的社会交互活动，考试则是主试者与被试者两种主体之间的社会交互活动；教育的职能是培养人才，开发智力资源、实现劳动力的再生产，考试的职能则是检测和甄选人才、开发人力资源、实现劳动力的有效利用。

所谓密切关联，就在于两种活动相互渗透、对象同一、目标一致、归宿相同：在教育活动中，考试是重要的组成部分，考试活动本身亦是一种特殊形式的教育；学校考试制度是国家考选制度的主要成分，又是国家教育制度的重要内容；教育与考试的对

①陈选善：《教育测验》，正中书局1933年版，第145页。

象都是人；两者均以社会对人才的需求标准为目标；服务于社会生产与社会生活，是考试和教育的共同归宿。因此，考试和教育都是社会发展的重要机制。

从社会大系统的角度考察，“人才培植为学校，而任使则属之各级机关”，“考试机关居中介之津，当关捩之地”，“教、考、用三者能密切配合，乃收人尽其才之功”。①一个社会，如果缺少教育，便必无所需之才，事业将难以为继；有教育而无考试，不仅培养目标难于期及、所育之才质量优劣不可辩，而且不能对人才作统筹有效之用；即使有健全的教育与考试，而任使不得法，仍不能收到良好的社会效益。无论哪个社会，只有教、考、用三者密切配合与协调运转，才能不乏人力资源和充分利用人力资源，并推动社会向前。

从学校教育系统看，考试处于从属地位，它受教育内外部规律的制约。但是，从属并不等于毫无独立性可言。学校教育系统的考试，同样也有相对的独立性，并能动地作用于学校教育的内部效率和外部效益。这就是方贤明先生所说的：“教育是国家百年树人的大计”，“应以保证最高限度地开发每一个人的才能为原则”，“更应审顾客观、主观的环境与社会的需要，为国家社会造就各种有用的人才”，而考试“实为完成此一目标的关键”。②

以上所述，仅是考试与教育关系之大要、事实上考试与教育有着多重主从和交叉关系，需有多方位的综合性系统研究。

本章拟在分析教育观、教学观与考试观三者关系的基础上，主要就学校考试与学校教育的关系问题，作一些具体的探讨。

---

①考试院考铨丛书指导委员会编：《中华民国考选制度》（台湾）中正书局1982年版，第2页。

②《世界高等教育改革动向》，（台湾）幼狮文化事业出版公司1973年版，第74页。

第一节　教育观、教学观与考试观

考试与教育，是因社会生产和社会生活对人才培养、甄选、任使的客观需要而产生的一对孪生子；考试观与教育观，又是反映社会人才需求和不同阶级人才意志之人才观派生的一对孪生子；而教学观又是教育观的派生物。一般说来，有什么样的人才观，就有什么样的考试观和教育观；有什么样的教育观，就有什么样的教学观。所不同的是：

其一，在不同的社会领域，教育观、考试观、教学观三者之间的关系并非固定不变。比如，在人类社会这个母系统中，考试与教育是两个互有关联而又彼此独立的子系统，各有其特殊的职能：教育依据社会人才需求育才；考试按照社会人才需求检测、甄选人才。尽管二者都以社会对人才的需求标准为目标，但考试处于育才与社会用才的中介地位，教育的"产品"合格与否？一般是以考试检测的结果作为判定的依据，在这个意义上，考试的承认亦即社会的承认。也就是说，考试对教育的调节作用客观存在。所以，就人类社会大系统的范围讲，考试观对教育观有一定的制约性。然而，在学校教育系统中，考试却处于从属地位，一切在学考试都应如实体现教学的任务，各类考试目标的确立必须分别以课程、教学、教育目标为据，因而考试观受制于教育观，而教学观又同时受到教育观和考试观两者的制约。

其二，在人类社会的不同历史阶段，教育观、考试观和教学观的内容与性质，也不是一成不变的。"实践、认识、再实践、再认识"，多次反复，步步深入，这是人类认识活动的基本规律。"不论在变革自然或变革社会的实践中，人们原定的思想、理论、计划、方案，毫无改变地实现出来的事，是很少的，"①

①《毛东泽选集》(合订本)，人民出版社1971年版，第270页。

随着客观现实的变化和实践的向前，人们面对新的任务，都要在新的实践中进行再认识，不断修正、发展、完善原有的思想和理论，逐步达到主观与客观统一、观念与存在一致，进而取得预期的效果。因此，生活在不同社会或不同历史阶段的人们，由于所处时代物质生活的生产方式不尽一样，社会的分工及其对人才种类与规格的需求不尽相同，他们都有一个再实践、再认识的问题，并在此基础上形成与其社会或时代人才需求相符合的人才观，借以指导人才的培养和选拔，使育、选之才适合社会发展的需要。既然人才观是随社会的发展和实践的向前而不断变化发展的，在不同的社会或历史阶段有着不同的性质和内容，那么，由人才观派生的教育观和考试观，以及从属于教育观的教学观，其性质和内容也理应伴随人类社会对人才需求的变化而变化、发展而发展。

以上表明，人才观、教育观、考试观和教学观，是随着人类社会的演进、人类实践的向前和认识的深化，而逐步形成的，并始终处在发展变化中（有关人才观、考试观产生、发展的演变过程，在第四章中已作叙述），但是，人才观、教育观、考试观三者之间的关系，并未因社会的变迁、时代的更替而改变，即教育观、考试观是人才观派生的，两者既是并列关系，又有相互制约的一面；教学观从属于教育观，同时也受到考试观的左右。对这种关系，我们还可以从传统人才观、现代人才观与教育观、考试观、教学观之间的相互联系及内在机制方面进一步探讨。

一、传统教育观、教学观与考试观

这里所说的传统教育观、教学观和考试观，是相对现代教育观、教学观和考试观而言的，并不是一种贬义，不能将它们与腐朽、落后保守简单等同，更不能将传统的与现代的教育观、教学观、考试观绝然分割开，或对立起来。因为世间任何事物都有一

个产生、形成、发展的演变过程，其发展的不同阶段都是前后关联的，常常是在前一阶段孕育着后一阶段的胚胎，而在后一阶段也往往还带有前一阶段的残余，一切新的因素，都是在旧的环境中逐渐萌芽、成长，才最终取代旧因素而成为事物的主流。传统的与现代的教育观、教学观和考试观，是就其存在的不同时代及性质、内容方面的变化而区分的。这是首先要说明的一点。

传统教育观、教学观和考试观，是传统人才观的具体体现，是封闭式自然经济的反映。在封闭式自然经济的社会里，为自给自足的生产方式所决定，人们的思维、态度等处于“定势”状态，以恪守祖宗之法、因循既定陈规为信条，对人才的要求是知书达理，虔诚上帝，忠君尊权，安分守已，进而形成了以知识蓄积为主要特征的传统人才观。由此派生的传统教育观则是：教师中心，唯书是从，传授为本，重在记忆，在此种观点支配下的传统教育，具有“崇尚书本，呆读死记与棍棒纪律，理论与实际分离，教育与劳动分离”①的特征。与此相适应的教学观便是：教师以传授现成文化知识、技艺为天职，学生以蓄积知识、掌握技艺为已任，其教学方法为注入灌输、死记硬背，重复书本之说，专与考试对口。所以，在知识蓄积型人才观居支配地位的时代，社会是以知识积蓄的多寡衡量和评价人才。传统的学校教育，因受传统教育观的制约，以及传统考试观的影响，一般是重知识灌输，轻智能发展；重死记硬背的书本知识，轻操作能力的培养；强调谨慎因循，不提倡开拓竞争；强调平均发展，不提倡发展个性，教学内容注重成果的记忆，教学程序刻板，学生只能如令行事，缺乏独立精神，虽吸收累积了成堆的知识，却失去了科学方法、科学态度与实践能力的有效训练。

传统考试观同样派生于知识蓄积型的传统人才观，它与传统

①中央教育行政学院编：《高等教育教学改革探讨》(内部发行)(上)，第70页。

教育观是孪生兄弟。在学校教育中，传统考试观还受制于唯书唯上的传统教育观。重知识、轻能力，重检测、轻督导，以记忆书本知识的数量与质量、掌握技艺的准确与熟练为标准评量、选拔人才，是传统考试观的主要特征。具有悠久历史的传统考试，就是在传统考试观与教育观的共同作用下，根据传统人才观的要求建立、发展起来的，其基本模式可概括为：内容照搬书本知识，命题靠少数人闭门造车，题型以论述题为主，测试方法起决定作用的仅笔试一种，评阅试卷全为手工方式，按评卷者个人对书本知识的理解主观判分，依据考生所得自然分数的高低定级选优，被试者以记背死知识的多寡、掌握技艺的准确与熟练程度取胜。可见，传统考试观一方面受到传统教育观的制约，另一方面，它又在客观上扶持了唯书唯上、注重知识技艺传授、强调记忆成果、轻视能力培养的传统教育观，而传统考试观对传统教学观的影响，也因此得到强化，以至造成教师因考施教、学生为考而学的严重局面，考试实际上起到了左右教学的作用。

由于传统教育观、教学观与考试观产生于自然经济时代，它必然要受到社会历史条件的制约，若用现代社会的观点衡量，其弊端是极为明显的。因此，对于传统教育和传统考试中那些不适应现实要求的陈腐东西理应扬弃。但是，在漫长的传统教育和传统考试中，也有许多珍贵的遗产，即使在现代教育和考试中，仍然具有实用价值。如传统教育注重知识传授的观点，今天还必须作为学校教育的基本指导思想，正如孙喜亭在《关于对教育的再认识问题》一文中所论述："从学校的根本任务看，学校首先（或主要）是传授知识的地方，学生在校期间要用主要精力和时间去掌握文化工具和学习基础理论，并取得智力的发展，也就是说学生要以主要精力和时间去接受前人积累下来的间接知识。否则，学校存在就失去了它的必要性。"又如传统考试以问答题作为测试的主要题型，应试者根据所问内容自由应答，这种试题在

检测学生能力方面，有其独具的优点：应试者可冷静深入地思考问题，自由阐述个人的思想与见解，充分展施自己的聪明才智；主试者通过测试结果，能了解应试者的知识水平和逻辑思维、发散性思维、创造性思维、文字表达等多方面的能力，有利于人才的发现、选拔以及个性的培养；大多数学科均可采用，具有比较广泛的适应性。此种试题不仅现实考试要继续沿用，就是我国全面实行标准化考试后，也将在试卷中占有一定的比重。再如“因材施教”、“学不躐等”、“教学相长”，以及作文考试，等等，这些传统的教学思想和考试方法，对提高现行教学质量、测试学生的综合能力来说，都是可资借鉴的。

二、现代教育观、教学观与考试观

恩格斯曾指出：“过去的资产阶级革命向大学要求的仅仅是律师，作为培养他们的政治活动家的最好的原料；而工人阶级的解放，除此之外还需要医生、工程师、化学家、农艺师及其他专门人材，因为问题在于不仅要掌管政治机器，而且要掌管全部社会生产”。①在这里，恩格斯从一个侧面揭示了社会发展与人才的关系，即在不同生产方式的社会，因其历史使命及经济、产业、技术、职业结构的不同，它对人才个体的素质与智能及人才整体结构的要求也不一样。或者说，社会的人才要求，以及反映这种需求的人才观，总是随社会形势与任务的变化而变化的。当自给自足的生产方式为现代生产方式取代后，社会的经济结构、产业结构、技术结构和职业结构等，都相应发生了深刻的变革，社会分工日趋细化，劳动过程的社会结合不断加强，体力在劳动技能要素中的地位大大削弱，而知识与智力则成为现代生产者劳动技能的基本要素。于是，社会对生产者的素质与智能提出了新的标

①《马克思恩格斯全集》第22卷，人民出版社1972年版，第487页。

准，要求现代生产者知识宽广，视野开阔，智能结构合理，敢于探索，勇于创新，富有开拓精神和良好的应变能力，进而迫使知识蓄积型的传统人才观让位于开拓创造型的现代人才观，并由此产生了现代教育观，即在系统传授科学文化知识的基础上，置智力开发和能力培养于首位，全面发展学生的智能，使受教育者在德、智、体、美、劳诸方面都能健康发展，具有高尚的道德情操、丰富的科学文化知识、熟练的劳动技能、实际工作的能力、开拓创新的精神和强健的体魄。

现代教育观的形成和发展，又促使教学摆脱了传统观念的束缚，改变了单传式的、“祖传秘方”式的、“填鸭式”的教学方法，扭转了教师中心、学生被动从命的局面，逐步确立了传授知识、培养能力、开发智力的现代教学观，其教学方法是启发诱导、引导发现、自我探索，在教学中坚持贯彻理论与实践结合的原则，充分发挥教师的主导作用，重视调动受教育者的能动性，引导学生成为学习的主人和发展的主体，启发诱导学生独立获取知识，“引发学生探讨心理的点燃”，激发创造的动机，促进个性才能的特殊发展，使学生能发表、能研究、能思考、能创造，学会学习并掌握科学知识的方法，把用脑与用手有机结合，为终生的学习和工作奠定坚实基础。现代教学观的主旨在于：它不把教学过程看作单纯传授与学习文化科学知识的过程，而是视为“促进学生全面发展的过程”，“要求教师在引导学生掌握知识的同时，全面发展学生的智力和体力，培养独立学习能力、学习兴趣和良好的学习习惯，以及从事创造性活动的能力”，①使学生对日新月异的现代生活及未来的工作具有更大的适应性。

考试观的发展与教育观的发展是并驾齐驱的，它也必须伴随人才观的演变而演变。现代人才观取代传统人才观后，它不仅要

①《中国大百科全书》（教育卷），中国大百科全书出版社1985年版，第153页。

求教育根据现代大生产的需要培养人才，而且要求考试按照现代大生产提出的人才标准检测、甄选人才，以便适应新技术革命的发展和社会生产部门对劳动者素质与智能要求的变化。这就不能没有与现代人才观相适应的新式考试。而考试的革新首先需有考试观念的变革。因此，现代人才观的产生，又导致重知轻能、单纯检测、不顾督导的传统考试观念的更新，以及现代考试观的形成。

现代考试观的基本观点是：全面检测基础知识，着重能力测试，以知识、智能结构是否合理为标准评量、选拔人才，督导教学双方实现教育目标。它要求考试如实反映社会需求，严格遵循教育、教学及人的身心发展规律，精心设计测试内容，慎重选择测试方法，按照严密的科学程序实施，有效控制误差，测试结果真实可靠，能准确反映被试者在知识、技能与智力方面的差异。通过对测试结果的分析，能为使用测试结果与利用反馈信息的部门或个人提供可靠依据。当今流行的标准化考试，就是根据现代考试理论，借助统计方法和电脑技术，严格按照系统的科学程序设计和实施，具有统一标准的测试活动，可谓现代考试的典型模式。现代考试因其理论科学、程序严密、手段先进、标准稳定、误差有控、结果准确、用途广泛，较之传统考试，它更能适应现代化社会对人才培养、甄选的要求。

现代考试和现代教育，既是现代社会的产物，又是推动现代社会发展的重要机制。它们互为依存，彼此配合，相互制约，为着同一个目标各以不同的方式服务于现代社会。

第二节　学校考试与学校教育的关系

要弄清学校考试与学校教育之间的关系，首先应对学校教育

的基本任务和学校考试的功用有一个比较明确的认识。

按照历史唯物主义的观点，学校教育是一种有目的、有计划、有组织的培养人的活动，是“使受教育者社会化”、“将知识形态的生产力转化为直接生产力”的一个过程。它的基本任务是：根据一定社会（阶级）的要求，按照一定的方向，采取适当的内容，借助有效的方法和手段，对受教育者进行系统科学文化知识和技能的传授，能力、智力、体力和思想品德的培养，使受教育者形成与其社会需要相适应的素质与智能，成为一定社会或阶级所需求的人。正由于学校教育是根据一定社会或阶级的要求培养人，所以学校教育的任务、内容以及性质和所起的作用，要受到社会政治、经济与科学文化发展状况的制约。学校教育究竟培养什么样的人，向受教育者传授哪些知识和技能，应使受教育者具有什么样的能力和思想品质，都取决于所处社会的历史条件。因此，在不同的社会或不同的国家，对学校教育的任务均有不同的规定和要求。比如，社会主义学校教育的基本任务就规定为，向学生传授系统的科学技术知识和基本技能，发展学生的智力、能力和体力，培养辩证唯物主义世界观和共产主义道德品质，使受教育者成为德、智、体、美、劳全面发展的社会主义劳动者。

在学校教育中，教学是实施教育的基本途径，学校教育各项任务的完成主要通过教学。而学校考试是教学活动中的有机组成部分，是教学过程的一个重要环节，它与其它教学环节一样，负有培养人的责任。又因其职能和功用的特殊，学校考试不仅本身理应体现学校教育的基本任务，而且是学校教育实现基本任务必不可少的措施。

有关学校考试在学校教育中的功用，本书在考试功能一章中已有阐述，不再细叙。总之，学校考试的功用有如下十二项：

（1）为决定新生取舍提供依据；

（2）为新生年级、班、组的安排提供依据；

（3）诊断教学问题；

（4）检测教学进度及其效果；

（5）督促教学；

（6）为决定学生升降、毕业提供依据；

（7）为教学改革与教育行政管理决策提供反馈信息；

（8）为评价学校教育质量提供依据；

（9）维持教学、教育标准；

（10）为家庭教育提供信息；

（11）预测学生将来的成就；

（12）为职业或就业指导提供依据。

从以上简述中，可知学校考试与学校教育的关系，亦即学校考试与学校教育内部和外部诸因素之间的相互联系、作用与影响。它与其它教学环节既密切关联又相对独立，既互相促进又彼此制约，各以不同方式发挥其在实现学校教育基本任务中的作用。就学校教育的整体而论，学校考试只是学校教育工作的一个方面，它的指导思想、目标、内容、形式、方法、手段、管理等，都必须服从学校教育人才培养的总目标，要受制于学校教育的内外部规律。然而，学校考试作为学校教育实现培养目标的重要手段，它又能动地作用于学校教育的其它方面。这突出反映在以下方面。

一、教育目标制约考试目标

教育目标和考试目标，是一定社会（阶级）对育才、选才要求的具体体现。培养什么规格的人才，按什么规格鉴别、甄选人才，其标准均为一定社会所提出，而不以教育者和考试者的主观意志为转移。因此，一定社会对人才素质与智能要求的标准，应是教育和考试确立目标的共同依据。这就是说，教育和考试作为人

类两种独立的社会实践活动，它们的目标（育才目标与选才目标）一致、标准同一，两种目标是相互独立的并列关系，互无统属。

学校考试目标与学校教育目标之间的关系则不然。尽管这两种目标也具有一致性，但两者之间的关系却不是并列，而是统属，即学校考试目标从属于学校教育目标。这种从属关系是学校考试本身在学校教育、教学活动中的地位所决定的。学校考试作为学校“教育流程三大要素”之一，它必须服从教育目标。而作为教学过程的重要环节，又必须服从教学目标和课程目标，如实反映其内容与要求。无论是为检查各门课程教学进度举行的考试，还是对各学科教学效果的检测，以及为鉴定某级教育是否达到规定培养目标所进行的考试，都要以教学大纲规定的相应目标为依据，分别确定检测的标准。即使属于社会考试范畴的各级招生考试，其测试标准，也要受到上下两级教育目标不同程度的制约。以我国一年一度的高校招生为例，它的测试标准的确立，在很大程度上要受制于中、高两级教育目标，既要反映高等教育培养目标的需求，又要基于中等教育培养目标的实际，做到需要与可能有机结合。倘若确立高校招生考试标准仅以高等教育目标的需求为据，而置中等教育的培养目标于不顾，便会导致中、高两级教育衔接不紧，影响人才培养的连贯性。相反，如果高校招生考试的目标完全以中等教育的培养目标为标准，不顾高等教育培养目标的需要，又将削弱高等教育的基础，影响高等教育目标的实现。近此年来，高校普遍反映新生适应能力差，“高分低能”的现象较为严重，有碍于教学质量的提高。出现这些问题，固然原因很多，但确立高校招生考试目标过于强调中学教育实际，与高校培养目标的要求距离太大，相应地降低了高校新生质量的检测标准，当是重要原因之一。这说明能否恰当处理高校招生考试目标与中、高两级教育目标的关系，将直接影响中、高两级教育

的衔接，以及高等学校的成才率。校外考试目标与学校教育目标的关系尚且如此，作为学校内部考试的目标，就更不能脱离教学大纲规定的目标或另立标准。否则，学校考试目标就不能保持与教学任务、培养目标，乃至学校教育方向的一致性，达到体现并督导实现学校教育目标的目的。

二、教学内容制约考试内容

在学校教育中，凡属检测教学效果的考试，如若离开"学校给学生传授的知识和技能，灌输的思想和观点，培养的习惯和行为"①等教学内容，要构成真实反映教学目标要求的测试内容，都将是"无米之炊"。可以说，学校教育的教学内容，是学校考试的物质基础。

人类教育发展的历史表明，学校考试的内容总是随学校教学内容的变化发展而变化发展的，奴隶社会时期产生了学校，逐步形成具有一定体系的教学内容，同时也就出现了反映这一时期教学内容的学校考试。如奴隶社会时期斯巴达的学校教育，它以体育和军事训练为主要教学内容，目的是培养效忠国家、性格残忍、具有镇压奴隶之军事才能的武士。当时为检验学校培养的人才是否符合统治阶级规定的武士标准，需经过严格而残酷的考试，其测试内容为：跣足行走、着单衣过严冬，夜间偷袭、捕杀奴隶、角力和球类竞赛等，在考试中基本上再现了教学内容，只是较之平日教学要求更高、标准更严。而在雅典商业贵族创办的文法、弦琴、体操等学校中，由于教学内容除体操、军事方面的训练外，还有读、写、算、音乐、文学、政治、哲学等方面的内容，所以学校考试的内容也有别于斯巴达：在文法、弦琴学校，对学生要进行字母、缀音、书法规则、诗歌、曲谱等内容的口头

①《中国大百科全书》(教育卷)，中国大百科全书出版社1985年版，第155页。

背诵测试，以及音乐、舞蹈、书写的演示测验；体育学校教学效果的检测内容为：角力、竞走、赛跑、跳跃、骑射、游泳比赛、掷铁饼、投标枪等。中国奴隶社会的教学内容与考试内容，又有别于西方，礼、乐、射、御、书、数“六艺”，是这一时期学校教学的主要内容，学校考试也同样是用这些内容检测学生的道德规范、征战技术、读写能力和数理知识。进入封建社会后，西方僧侣封建主的教育，以文法、修辞、辩证法、算术、几何、天文、音乐“七艺”为主要教学内容，而世俗封建主的教育，则以骑马、游泳、投枪、剑击、打猎、下棋、吟诗“七技”为主要教学内容，中国封建社会的主要教学内容则为，“五经”（《诗经》、《书经》、《易经》、《礼记》、《春秋》）和“四书”（《论语》、《孟子》、《大学》、《中庸》），有时还兼设医学、数学等实用科学方面的内容。在此期间，不论是西方的“教士”、“骑士”考试，还是中国的“选士”、“科举”考试，所考内容无不来源于上述教学内容。到了近现代，由于社会生产和科学技术的迅速发展，学校的教学内容日益丰富，学校考试的内容也随之更新，并日趋综合化、系统化。现阶段，一份学科考试的试卷，通常包括所测学科教学目标要求学生掌握的基本原理、概念与规律性知识，以及学生必须达到的不同层次认知水平方面的内容，以便全面、系统地检测教学效果。但是，一份学科考试试卷内容的多寡，完全取决于学科的教学内容是否翔实，没有充实的教学内容，考试内容就不可能有丰富的素材来源，况且学科考试必须体现教学的基本任务。所以，学校各种检测教学效果的考试，其测试内容受到教学内容的制约，是必然的，也是正常的现象。

三、学校考试影响学校教育的效率和效益

学校考试既受到诸如教学目标、教学内容、教育思想、教育

目的等学校教育内部因素的制约，同时又给学校教育以多方面的强热影响。它对教学的进程与质量，学生知识、智能结构的形成，教学内容、方法、手段的改进，教育水准的升降，乃至所育人才的社会适应能力，均能起到调节控制的作用，或扶持、推动，或压抑、阻碍，直接影响学校教育的效率和效益。这可从学校教育内、外两方面考察。

从学校教育内部看，学校考试主要影响的效率，即影响学校教育质量的提高。其表现是：

1．左右教学内容和教学方法。

马克思曾以大量事实论证："不是意识决定生活，而是生活决定意识"。尽管各级学校都有规定的教学大纲，本应以此作为组织教学的依据、检查教学进度和衡量教学质量的标准。但由于种种社会原因，考试成绩在现实生活中的作用举足轻重，无论是学生的升级、留级、毕业、获得各种荣誉称号，还是教师的晋级、调资，乃至学校的社会声誉，都与学生的考试成绩密切关联。因此，在教学工作中，考试具有一定的支配作用，教学双方一般是根据考试的"行情"选择教学的内容和方法，客观上存在着因考施教、为考而学的现象。这就是人们通常所讲的考试的"导标"作用。

考试左右教学内容与方法，既不是始于现实社会，也不只是在中国，古今中外莫不如此，只是这种作用有二重性而已，即对教学双方的影响有良与不良之分。要人为的消除考试对教学的影响是愚蠢的，也是办不到的，这是考试固有的功能。因此，问题的实质，不在于回避或一味批判考试对教学内容和方法的影响，而是要深入研究两者之间的关系，把握考试左右教学内容与方法的规律，科学设计考试的形式、内容与方法，使之服从教学大纲的需要，正确发挥"导标"功能，给师生的教与学以积极良好的影响，实现教、学、考有机统一、协调一致。

另一方面，考试的反馈信息是学校研究教学的重要依据之一，对教学内容、教学方法以及教育思想与手段的改进，能起到促进或阻碍的作用。学校每进行一次考试，不单是为了获得反映教学成就大小的一个标志——考试分数，更重要的是，通过对考试分数及其它数据资料的综合分析，全面了解教学中的各种情况，譬如：教学大纲的要求是否适合教育对象的实际？教材内容是否科学？教学进度是否适度？讲授方法是否得当？学生学习的难点何在？影响教学效果主要因素是什么？等等，以准确掌握教学情况，为教学选择恰当途径取得可靠依据。因而，考试的反馈信息对学校改进教学至关重要。若信息准确、客观并能科学利用，便可成为教改的动力，有助于教学双方避短扬长、除弊兴利，促进教学工作健康发展。若考试提供的信息缺乏真实性，或与教学的实际情况相反，又将成为教学的阻力，妨碍教学中健康因素的发展，造成教学因循陈规，缺乏主动适应社会要求的能力。在这方面我国历时1300年的科举考试，有极为深刻的教训，也有可资借鉴的经验，归结起来就是一条：考试本身科学与否，是决定影响好坏的根本所在。没有科学的考试，就得不到准确可靠的考试结果。而考试结果不真实，即使有先进的统计分析手段，也不可能为教学或教育行政管理部门提供客观反映教学情况的信息，进而产生健康的影响。所以，要避免考试信息对改进教学的消极作用，首先必须治本，即根除考试本身的流弊。

2．影响学校教育的成才率。

人才总是通过与现实世界相互联系、相互作用而逐步成长起来的。在人才成长的过程中，虽然影响人才的因素很多，除先天资质、生理发展的成熟度、兴趣、情感、动机、意志力外，还有教育、环境、营养和社会实践等，但总起来看，不外乎两大因素：内部因素和外部因素。每一个人才的成长过程，都是这两大因素互为条件、相互作用、不断调节的过程。从学校育才的角度

看，考试对成才的影响属外部因素。它能借助考试的内容、方法，主试者的言行、应试环境及考试结果，激发和诱导被试者思想、动机和意志的产生与发展，促使被试者向着规定的培养目标进取。其原因在于：学校考试是教学活动的重要组成部分，考试的内容要体现学校教育的基本任务，反映学校对人才培养的要求，所以考试具有教育性、目的性和竞争性，能够强化学习动机和成才意识，形成追求培养目标的自觉行为。但这并不是说，凡学校都有这种效应。考试是主试者与被试者之间的双边活动，而考试的教育性、目的性和竞争性，又是在考试主体（学生）与客体（考试内容）的相互关系及交互作用中得以实现的。要使考试取得促进学生成才的效果，需具备两项条件：一是主试者的言行、应试环境、测试方法要适合被试者的身心发展规律，使被试者以良好的心理状态主动积极地接受考试；二是恰当处理考试主体与客体之间的关系，使主体面向的客体真实反映学校教育的基本任务与要求，并为考试主体所理解，使考试主体的意志从属于这个目的，"成为'理想的意图'，并且通过这种形态变成'理想的力量'"，[1]才能使被试者产生追求培养目标的行为，展开正确思想指导下的竞争，积极主动地配合学校教育，把自己转化为德、智、体、美、劳全面发展的人。

在学校教育的外部关系方面，学校考试影响学校教育的社会效用和人才培养的经济效益。

学校教育为其性质、任务所决定，它的存在和发展务必注重两点：在内部效率方面，应力求在发展数量的同时，保持教育质量的稳步提高；在外部效益方面，要设法使所培养的人才能够适应社会发展的需求，产生良好的社会效用。要做到这一点，从一个国家来说，需对以培养人才为要务的各级教育分别作出统一的

[1]《马克思恩格斯选集》第4卷，人民出版社1972年版，第228页。

要求，并对各级教育培养的人才有明确的合格标准，以此作为检验人才质量的尺度。而对一所学校来说，一切教育措施都应服从规定的培养目标，有利于学校教育按照社会和国家的要求完成其使命。历史的经验证明，在维持学校教育人才培养标准的各项措施中，考试的作用实不可轻视。

学校考试不仅是督导学校教育实现培养目标的手段，而且是检验学校教育是否达到规定标准的工具。它对学校培养的人才能否适应社会的需要有控制作用，直接影响学校教育的效益。这种控制作用及其影响，一般是学校阶段性考试和毕业考试所致。

1．学校阶段性考试调节控制人才成长的“营养”，影响教育投资的经济效益。

一个人在求学期间培养起来的兴趣和爱好，是将来发奋拼搏、创造性学习和工作的有效能源；求学期间形成的知识、智能结构，是一个人成才的坚实基础。而求学期间个体心理的发展和知识、智能结构的形成，主要仰赖良好的教育，同时与学校考试的影响密切相关。我国近代著名教育家陶行知先生在总结自己多年育才的经验时指出：“小孩的体力与心理都要适当的营养，有了适当的营养，才能发生高度的创造力，否则创造力就会被削弱，甚而至于夭折。”这里所强调的“营养”，包括物质与精神两个方面的营养，可谓新生一代成长的物质基础。作为在校学生，他们心理发展的主要“营养”，乃是全面系统的基础知识和基本技能。没有这种“营养”，学校就造就不出全面发展并具有高度创造力的人才。这就涉及到教师给学生什么“营养”、如何给“营养”，以及学生吸收什么“营养”、如何吸收“营养”的问题。恰恰在这个问题上，学校阶段性考试有调节控制作用。

学校阶段性考试的设计与组织（如单元测验、期中考试）常有两种不同的目的：一种是因教设考，一种是为考设考。由于目的不同，产生的效果也全然相异。

因教设考的主旨，是将考试与培养目标对口，让考试服务于教学，对人才成长的所需“营养”以合理调节。在此种考试思想指导下，学校在不同教学阶段举行的考试，其节拍与教学一致，目标与教学大纲相对应教学阶段的要求相符，内容包括教学大纲规定的难点、重点和必须掌握的原理、概念与基础知识，全面检测教学双方的效果。通过考试，对学生学业的增进速度，教师阶段教学的成效，整个教学的变化趋向，以及影响教学质量的主要问题等，能获得比较完整的反馈信息。就此进行分析研究，诊断教学中的弊端，查明学生对“营养”的所需，以及教师供给“营养”的数量、质量与方式，无疑有助于调解“营养”的供需矛盾，使学生智能发展方面的缺陷及时得到弥补，不至出现饥饱无度、“营养”失调的现象。这样的考试，有利于教学双方的密切配合，促进教师系统讲授教学内容，引导学生全面掌握基本理论和科技知识，形成合理的知识、智能结构，也有利于学校教育提高人才培养的经济效益。

因考设考则不然，它的主旨是将学校阶段性考试与校外招生考试的目标对口，让教学为考试服务，以校外招生考试的需要为标准供给学生“营养”。受此种考试思想的支配，学校教学以适应校外招生考试为目的，以升学率这一孤立的指标作为评判教学、教育质量的标准。平日所组织的大大小小的考试，完全模拟校外招生考试的形式、内容和方法进行，师生专注于校外招生考试有关内容的灌输与记背。教师给学生的不是自我“谋生”的本领，仅是维持短时冲刺的“干粮”；不是怎样过河的若干方法，而是辛辛苦苦地背着学生过河，当学生从教师背上下地后，吃完“干粮”，就丧失了自我行走的能力。也正是这种考试导致学生“营养”失调、知识缺漏，削弱了学生身心健康成长的物质基础，为学生的继续深造和创造能力的发展设置了难以逾越的障碍，无力成为适应社会发展需要的开拓创造型人才，从而造成教

育投资的浪费。

2．学校毕业考试影响学校教育的"产品"质量及社会效用。

学校毕业考试，在各级学校教育中位居津要，对学校培养人才的质量负有把关的重任。这一把关措施是否可靠有效，不仅关系到高一级教育的基石是否牢靠，并妨碍整个民族文化素质的提高，影响学校教育的"产品"质量及社会效用。新中国成立以来，千百万中、高两级毕业生成长的历史与工作现状，以及建国后不同时期中、高两级教育的质量与毕业考试的质量及毕业生社会效用对比研究的结果表明：各个时期以上三者间的关系呈线性状。据一所师范大学的毕业生调查资料载，该校对近30年分配到某自治州的千余名毕业生进行了全面的调查。分配去的毕业生，除少数分布在各级党政群机关外，其余800多名均在大、中、小学任教，其中"文革"前毕业的占45%，"文革"中毕业的占30%，"文革"后毕业的占25%。调查结果表明，不同时期毕业生的工作能力与成效，与各自所处时期中、高两级教育的教学情况和毕业考试的情况，有着极为密切的关系。"文革"前，由于学校教学比较正规，学生所学知识全面系统，基础知识宽厚扎实，智能结构比较合理，毕业考试标准明确、把关严格，所以95%的毕业生，都能胜任所学专业的教学工作，并成为各级各类学校教学的主力军，其中不少人因教学所需，或由文转理，或文、理兼教。"文革"中毕业的少数后起之秀，大都是"文革"前初、高中基础比较扎实、当时毕业考试达标的人。其余部分是"文革"期间中、高两级学校毕业生，因考试制度被废止，没有严格的毕业把关措施，缺少统一的毕业标准，不论学习成绩优劣，学习期满都可毕业。与此相适应，他们的工作能力及成效相互间参差不齐，与前一时期的毕业生相比，社会效用有明显差异。"文革"后的毕业生，虽然考试制度得以恢复，但因校外招生考试的影响，中学阶段教学偏科，且不系统，学生所学知识残缺不全，又无严格

的毕业标准，凡在校就读的学生，只要未犯严重错误，都能得到中学毕业文凭。进入高校后多为单科独进，知识的横向联系未能足够重视，至于高校毕业考试，不管各学科考试目标是否体现了教学目标要求的标准，达到60分便可太平过关。基于上因，这一时期的毕业生到教学岗位后，限于知识面窄和运用所学知识解决实际问题的能力差，虽能基本胜任所学专业的教学工作，但可文、理转教或兼教者极少。这说明，健全的毕业考试制度，是各级教育把好"产品"质量关，强化社会效用的必要措施。

第三节 学校考试的范畴、类别、原则与方法

人类社会的考试，因社会分工对从业人员素质与智能的要求的多样性，而分化为众多的类别，以及同一类别的若干层次，关系极为繁复，大致可分为社会和学校两大考试系统。学校考试属于学校考试系统的一个组成部分。它的范畴、类别、原则与方法，应依据学校考试的任务与目的而确立，以区别于社会考试系统和学校考试系统的其它考试。

一、学校考试的范畴与类别

学校考试的范畴及其所含考试的类别究竟如何确定？现阶段人们的看法不尽一致。一种观点认为，学校考试与学校考试系统是同一概念，两者的范畴与类别是相同的。我们的观点是，学校考试与学校考试系统是从属关系，两者的范畴及其所属考试的类别有显著区别。学校考试系统的范畴，除学校考试外，还包括由学校外部主持实施、对象和内容属于学校教育范围的考试，以及由教育外部其他管理部门决定而在学校内部实施的考试，诸如各类各级学校教育的招生考试，评估学校教育质量的考试，出国留

学考试，各种竞赛考试，社会其他部门在学生中挑选特殊人才的选拔考试，教职员工的招聘、晋职、提级考试等。而学校考试的范畴，系指学校内部根据培养目标的要求，自行编制并主持实施的各种考试。安置新生、诊断教学问题、了解教学进度、检测学生学业成绩和教学效果，是学校考试的本职任务，目的在于掌握教学情况、改进教学与督导学校培养目标的实现。因此，学校考试与学校考试系统不是同一概念，二者不能交互使用。

学校考试的类别的划分多种多样，一般是随划分的角度而变化，常见的有如下几种分法：从考试时序的角度，分为学前考试、平时测验、期中考试、学期考试、学年考试和毕业考试；从考试方法的角度，分为笔试（开卷考试、闭卷考试）、面试（口试、操作演示考试）、个别考试和团体考试；从考试目的的角度，分为诊断考试、成就考试和合格考试；从考试内容的角度，分为单科考试和综合考试。在国家教育行政机关制定的学生学籍管理章程中，通用的是时序划分法，一般将学校考试分为平时测验、期中考试、学期考试、学年考试和毕业考试五种。各级学校大都以学期考试、学年考试、毕业考试的成绩，分别作为决定学生升留级、颁发毕业证书、授予学位证书和毕业后分配工作的主要依据。

二、学校考试的原则与方法

学校考试是检查学生学业成绩和教学效果的重要方法，也是实现学校培养目标的基本途径。它对"稳定教学秩序"，健全教学制度，改进教学的指导与管理，加强学校教育与家庭教育的配合，形成良好的校风，提高学校教育的效率与效益，都有十分重要的作用。切不可忽略的是学校考试的作用同样存在二重性，要想收到促进学校培养目标实现的良好效果，就必须在正确考试思想的指导下，严格按照学校考试的原则选择考试内容与方法。

（一）因教施考

尽管各级学校的培养目标、学科、门类、专业设置不同，教学内容、教材性质及课程的特点各具，对考试的具体要求也互有区别，但是各种不同形式与内容的学校考试，其基本任务无一不是根据教学大纲的有关要求或教学的需要，去诊断教学情况，检测教学效果，督促教学上进，维持教学质量的标准。因此，学校考试应坚持因教施考、以考促教的总原则，即检测教学效果的标准，敦促教学的要求，诊断教学问题、测度教学成效的内容等，应以教学大纲为准绳，既不能根据教师个人的意愿确立考试的内容为标准，也不能按照学生的实际情况随意降低或拔高教学质量的检测标准，更不能置教学大纲的要求于不顾，以适应校外招生考试的需要为原则设置学校内部考试，以至本末倒置，让教学服从于考试。这就是说，因教施考，绝不是按教师和学生的主观意志施考，而是要学校考试体现学校教育的基本任务，如实反映教学大纲的要求，通过考试，强化教学双方追求教学目标和培养目标的意识，促使师生自觉贯彻落实教学大纲的要求，努力实现大纲规定的教学目标。

（二）全面检测

这一原则的主旨是，学校考试应全面体现教学大纲对教与学的要求，既要检测学生的基本知识、基本理论和基本技能，又要检测学生运用所学知识、理论、技能分析或解决实际问题的能力。恰当处理下列考试内容设计的三大关系，乃是学校考试全面检测原则的基本要求。

1．重点与一般的关系。

一本结构合理的教材，其内容总是重点与一般有机结合，主干与枝叶纵横兼顾。一门学科的课程设置，其编排方式都是由浅入深，从易到难。同样，一门学科的教学大纲也是深度与广度的有机统一，它对师生传授和掌握教学内容的要求，在横向上与教

材结构相适应，突出重点，顾及全面，要求点面结合。在纵向上与课程编排相一致，从低到高，逐级上升，要求循序渐进。既然教材内容是学校考试内容的来源，教学大纲是学校考试命题要求的依据，那么，学校考试的内容，也就理应遵循点面结合、纵横兼顾的原则，正确处理重点与一般内容的关系。所谓重点，即是教学大纲要求学生必须掌握的关键性内容，诸如基本原理、基本概念、基本技能和一些带规律性的知识，亦即一门学科的主干成份。所谓一般，便是一门学科的辅助成份，亦即通常所讲的"面"。然而，一份试卷的容量是有限的，它不可能包罗一门学科教材的全部内容，只能择其要者而试之。这就需要在考前认真研究教学大纲的要求，吃透教材内容，以便考试内容的设计能重点突出，兼顾全面。一般说来，学校考试的内容应以点带面，将教学大纲规定的重点作为主体内容，着重考核学生对"三基"知识的掌握及其应用。同时，又要以面扶点，尽可能扩大测试的覆盖面，凡与各章节教学重点、难点相关密切的内容，都应有选择地纳入测试内容。重点内容的题型宜选用容量较大的综合性试题，一般内容的试题则宜小型多样。因此，一张试卷内容的比重点大于一般，而题量的比例当一般大于重点。在考试中，重点内容和一般内容的比例及所用题型是否恰当，不仅影响考试对教学效果的全面检测，而且影响到学生能否形成合理的知识结构，如果设计的考试内容只有重点没有一般，就"不能引导学生抓住主要内容；但如果只有点没有面，则学生往往只孤立地去抓所谓重点而不重视全面系统的复习"，①导致知识结构上的缺陷。

2．知识与能力的关系。

传授和掌握系统的文化科学知识，当是教学双方的主要任

①潘懋元主编：《高等教育学》（上），人民教育出版社和福建教育出版社1984年版，第304页。

务。但传授知识并不是学校教育的全部任务，更不是终极目的。知识的传授和积累，只是为学生形成一定社会所要求的素质与智能提供了“营养”，尚需采取各种有效措施培养和发展学生的智力、能力与体力，才能使之成为具有与其社会要求相适应的素质与智能的人才。因此，学校考试要全面检测教学效果，敦促教学全面发展，就应“贯彻知识积累与智力、能力发展相结合的原则”，既考知识又考能力。各学科知识考核的范围，以及认知能力考核的不同级别，当根据教学大纲的要求确立。至于知识与能力在考试内容中所占的比重，则应根据具体测试对象而定，要适合教育、教学和学生身心发展的规律。一般来说，初等学校的考试，应侧重考核掌握知识的数量与质量，技能的准确与熟练程度。能力方面重在记忆、理解两级认知水平的考查，试题以单项题为主，综合性试题尽量少用。中等学校的考试，当知识与能力并重，既要考查学生对学科基本原理、概念、技能掌握的准确与牢固程度，也要考查驾驭和综合应用所学知识的能力，其认知水平的考查，着重于记忆、理解、应用、分析四级，兼以综合、判断能力的考查。测试试题单项与综合两类并用，但综合性试题的比例不宜过大，约占试题总量的40%左右即可。高等院校的考试，必须坚持“既考知识又考能力，既考深入理解又考综合应用”的原则，并以能力的考核为重。论述题、判断题、论证题及综合性应用题，在试卷中要占有较大的比例，一般在试题总量的60%以上，所含内容应是测试内容的主体，但又要注重知识考查的面。否则，考试不利于促进学生全面打基础，进而影响能力的培养和发展。

3．理论与实践的关系。

学校考试既要考书本知识的掌握，又要考实际应用的能力。这是因为，学生自小学到大学，乃至研究生的求学过程中，主要是接受书本知识，为深化学生对基础理论的理解，促使学生更好

地掌握学科范围内的技能与技巧，达到学以致用的目的，教学过程中都要安排一定量的实践性教学活动，如社会实践、教学实习、科学实验、工程设计等。所以，学校考试除检测课堂教学的效果外，还必须注重教学实践环节效果的考查，应视之为全面检测的重要组成部分。尤其是实践性较强的学科，更要重视实践能力的考核。若忽略教学实践环节效果的考查，只注重课堂教学效果的检测，不仅不能全面检测教学效果，而且难以使教学任务的要求对学生的实践性活动产生主导作用，以及使学生的实践对实现教学目标真正具有从属意义。

（三）节拍一致

学校考试伴随教学过程的始终，具有检测教学效果，促进教学发展，以至左右教学方向的作用。因此，学校考试必须服从教学目标的要求，各种考试的设置应与教学的节拍保持一致，即以不同阶段教学的内容与要求设计考试的内容与目的，按照教学计划的规定制订考试计划，确施考的时间和次数，切忌随意增减考试的次数，或任意调整施考时间，甚至走极端，要么考试成灾，要么一次不考。这样做必将打乱教学的正常节奏，降低人才培养的效率。教学是一个连续性的渐进过程，学校培养人才的效率是随教学阶段递进的，这是教学的基本规律。作为检测阶段教学成效（或培养人才的阶段递进效率）的考试，只有与教学的节拍保持一致，将考试设在每一教学阶段的结束之际，并如实反映相应阶段教学的内容和要求，才能起到检前促后或前承后启的作用。

（四）方法恰当

学校考试如同其他考试一样，命题与测试对考试的成败有决定性意义。因为学校考试无非是根据一定的目的，用一定内容的试题或项目，通过一定的方式，去测量应试者的素质、知识、技能或能力，并以测得结果作为甄别应试者被测方面差异的依据。所以，一份考卷、一组测试项目，也就是测量应试者的一把量

尺。测量结果的可靠性和有效性，主要取决于量尺的准确性和施测的规范性，而量尺本身的精确度，在很大程度上又决定于命题方式。

在命题方式方面，现阶段中外学校考试常用的有两种：一种是集体命题，其特点是开放性、经常性；另一种是个体命题，其特点是封闭性、突击性。从命题程序及所命试题的实际效果看，两种命题方式也有明显的差异。

集体命题一般有一个相对稳定的命题班子，参与命题的人员通常是学科专家，对本学科的教学情况有深入了解，精通学科教材内容和教学大纲的要求，并熟知命题的基本理论与原则，有较高的命题技能、技巧，命题过程中能彼此配合、互补长短、协调一致，严格按照科学程序编制试题，即根据教学目标和考试目的的要求，共同拟定试题编制计划，作为指导命题人员编制试题的实施蓝图。命题人员分头编制的试题，要经过命题组多次集体审查和修改，再经预测检验，确系可靠有效的试题才用于正式测试。由于集体命题能较好地克服命题人员各自主观认识的缺陷，充分发挥集体的智慧，加之命题的周期长，可对编制的试题进行反复推敲、筛选，并通过预测作客观检测。所以，试题有较好的难度和区分度，考试标准比较稳定，同类考试多次测试的结果基本一致，很少出现分数升值或贬值的现象，能较为准确地反映被试者受测方面的真实水平。

个体命题，又可分为任课教师自行命题和指定某一教师命题两种方式，但命题特点、程序与试题的实际效果并无两样。个人命题出于保密需要，一般是命题者根据个人对教学大纲、考试目的、教材内容的理解，凭个人经验闭门搞短期突击，既无统一制定的编题计划作指导，又缺少试测和试题筛选的环节，试题完全由命题者个人经验主观决断。因此，试卷的难易程度及可靠性和有效性，命题者不能把握，考试标准往往是随命题人员的水平而

上下波动、左右偏倒，各次考试分数缺乏稳定的值，即不能准确体现教学大纲的要求，也反映不出应试者真实的学识、学能水平。

相比之下，集体命题方式在维护考试的可靠性、有效性方面，较之个体命题方式更有保证。但不能因此得出学校考试的命题只能采用集体方式的结论。命题方式应根据考试目的的需要抉择。在学校考试中，凡属阶段性、总结性和安置性的考试，如期中考试、学期考试、学年考试、毕业考试以及为学生分流、新生分班、分组举行的考试，命题宜用集体方式。至于各科教师为了解教学进度、诊断教学问题而举行的考试，目的在于教师本人调整教学内容、改进教学方法，此类考试由任课教师自行命题，就比集体命题更切合实际。

学校考试要取得良好效果，除恰当选用命题方式、保证量尺的准确性外，还应审慎选择考试方法。

现行学校考试的方法有笔试、面试两种。笔试又分为开卷考试与闭卷考试。面试包括口试、操作、演示、答辩等多种形式。考试方法是考试设计首先要明确的问题，它关系到编制什么样的试题或项目，以及采用什么形式施考。再者，不同的考试方法有不同的适应范围和效用，测试效果各有利弊。

笔试是学校考试中使用最普遍的考试方法，考生书面回答问题，作答形式分自由应答（主观性考试）和固定应答（客观性考试），多用于基本原理、基本概念、基础知识及记忆性、综合性、论证性、判断性、创造性项目的团体测试。采用笔试施测，考生能够比较冷静地思考问题，自由表达自己的观点，既可测量考生掌握知识的数量与质量，也可了解考生的逻辑思维能力、文字表达能力和发散性思维能力，并有利于阅卷者对考生成绩作出较准确的比较和评价。但笔试要受到文字与时间的双重限制，测试内容不及口试灵活，覆盖面相对口试较小，主观性试题的判分

标准也不易掌握。

面试是一种直观性的测试，通常采取对话、操作、演示三种形式检测学生的学业成绩，多用于口头表达能力、掌握技艺的准确熟练程度、掌握知识的深度与广度、理论联系实际的能力、思维的敏捷性等项目的个体测试。测试内容主试者可根据考生作答情况加深加宽，或降低难度，缩小范围，考生不受文字限制，尤其是研究生的论文答辩，考生可充分运用与测试内容有关的知识进行自由辩驳，大胆阐述自己的观点。面试不仅能比较全面深入地检测考生的学习质量，并有利于促进个性的发展和特殊才能的了解，它与笔试有互补作用。但是，面试也有工作量大，效率低，评分误差难于控制，考生答题的随机性大等缺点。

基于以上原因，学校所进行的各种考试，究竟采用哪种方法适宜，必须根据不同学科、课程、教材的性质与特点以及考试的目的要求，进行慎重选择。

学校考试是获取学生升级、留级、退学、毕业、授予学位依据的主要形式，也是督导学校实现培养目标的重要手段。因而，一切学校考试应遵循教育、教学和考生身心发展的客观规律。确定考试目的、设计考试内容、选择考试方法、使用考试结果，都要有利于教学的改进和质量的提高，有利于学生身心的全面发展。

第十一章

考试与人事管理

科学的考试是人事管理部门有效开发与合理利用社会人力资源的重要途径，而科学的人事管理又是考试结果由理论价值转化为社会价值的基本保证。

社会系统（可以是一个国家、一个地区，也可以是一个行业、一个部门或一个单位）中的人事管理部门（或子系统）的工作，是开发和利用本系统的人力资源，对本系统的人员从培养、甄别、选拔到使用进行全面的管理。由于人是世界上最可宝贵的财富，是推动社会发展的基本动力，所以，人事管理工作，是社会系统各项工作中处于特殊地位、具有特殊意义的工作。

我国春秋时期的人才大家管仲，他在论述人才的使用时曾说过："成器不用。"①其意思是说，即使是人才，不经过检测也不能使用。十六国前秦时的高泰，在回答苻坚（前秦的皇帝）关于治国之本的问题时指出："治本在得人，得人在慎举，慎举在核真。"所谓"核真"，用今天的话说，就是考核（所荐举之人的）真实情况。我国历史上选拔人才、选拔官吏的几种最主要的形式——原始贤能制、禅位制、军功制、科举制、考绩制和保举

①《管子·七法》。

制等所共有的特点，就是对人须经甄别、测度后再决定选录与否。现代社会系统的人事管理工作中，对人员（这里的“人员”是广义的，相当于“人力资源”）的考评，已成为一门科学——“人员测评工程”或“人员测评学”。由此可见，考试与人事管理之间的关系，不但源远流长，而且愈来愈密不可分。所以，认真地研究考试与人事管理之间的关系，以及人事管理考试的组织、原则、职能、范畴和类别等，对于提高社会系统人事管理工作的科学化程度，使社会系统中的各种人员各得其所、各尽其能等，都将大有裨益。

第一节 考试与人事管理的关系

有用人问题，就有相应的考试。过去是这样，现在是如此，将来(只要社会还存在分工)也不例外。随着社会科学化程度的提高，考试不仅是人事管理工作的一种工具和手段，而且还是人事管理系统一个有机的组成部分。

一、人事管理的社会功能

为了研究考试与人事管理之间的关系，有必要先弄清两者的社会功能。关于考试的功能本书第三章已作了详尽而系统的讨论。这里仅就人事管理的社会功能作一简略的介绍。

（一）科学的人事管理，能充分实现人的价值

一般说来，人的价值的大小与其自身才干的高低有关。但在人们自身才干已定的情况下，能否充分实现人的价值，则取决于社会系统对人的测度、甄别是否准确、使用是否恰当，亦即管理是否科学。

（二）科学的人事管理，可以使人员的使用最优化

作为组成社会系统的人，除有其共同的基本特点外，更重要的是各有其独特之处。因为人的素质与职能的个别差异客观存在，世界上不可能有两个在道德品质、思想方法、知识、能力、爱好、特长、体形、体质、体能等方面完全一样的人。这正是社会系统成员的个性所在。这种个性，既是本书第四章中关于人才分类的重要依据，也是社会系统中人事管理部门对人员实行管理的重要依据。量才用人并进行动态的追综管理，能使全体成员各尽其才、才尽其用，实现人员使用的最优化。

（三）科学的人事管理，能开发潜在的人力资源

所谓潜在的人力资源包括两个方面：一是发现新的人才；二是开发已有人员尚未发挥出来的才干。近几年我国出现的人才流动、科技人员兼职、退（离）休干部继续发挥“余热”等，就属于这种情况。

（四）科学的人事管理，能够减少以至避免人才浪费

人才浪费是最大的也是最不能容忍的浪费。目前我国人才浪费的现象还比较严重，积压人才、压制人才和摧残人才的情况，在整个社会系统中还不同程度地存在。这正是我国人事管理制度必须改革的重要原因之一。科学的人事管理，则可以减少以至避免人才的浪费。

从考试的行政功能和社会系统人事管理子系统的功能中，我们不难看出，考试与人事管理之间具有非常密切的关系。

二、考试在现代人事管理中的作用

（一）考试是现代人事管理的起点

在现代社会中，各企事业单位无论是招收新工人还是招聘新的专业干部，都要先经过考试、考核或考察，然后再视其结果决定是否录用。从这个意义上讲，考试是现代人事管理工作的起点。

（二）考试是现代人事管理的基础和手段

人事管理工作是一个过程，即对各类人员的使用情况根据一定的要求、目的不断地实施控制的过程。在这个过程中，一方面要不断地开发新的人力资源，以满足本社会系统正常运转和发展的需要；另一方面，要对各类在业人员进行追踪考察，以便对他们的道德品质、思想状态、知识能力、身体素质等方面的情况，作出与其当时实际情况相一致的判断。并以此为依据，对这些人员不断地实施管理，或继任现职工作，或调换工作职位。人事管理的过程，实质上是一个不断开发新的人力资源和对已有人员实施控制的过程。然而，无论是开发新的人力资源，还是对已有人员的追踪管理，都有一个了解人的问题。因为唯有“知人”才可能“善任”。而要了解人，就离不开考试，定性的或定量的、静态的或动态的、专门的或综合的，等等。所以考试是现代人事管理的重要基础。

现代社会不但知识密集程度愈来愈高，而且竞争也愈来愈激烈，为了适应这种新的情况，人力资源的开发已由过去的“体力型”演变为“智能型”。在现代社会系统中，任何人事管理工作的最终目的，无一不是为了使其系统在竞争中获胜。而对于人事管理工作来说，能实现这一目的的唯一途径，就是尽可能有效开发和合理利用本系统的人力资源。在社会系统成员智能化程度不断提高的情况下，要准确全面地了解人愈来愈重要，也愈来愈难。凭直观的印象或简单的考察已不能达到全面了解人的目的，必须借助科学的考试手段，才能对不同人员德、学、才、识、体方面的实际水平有较为准确客观的了解。甚至可以说，考试是现代社会人事管理中必不可少的条件。

科学的考试，对于现代社会系统中的人员来说，还具有另一种极其重要的作用，即通过考试可以正确地认识自己，了解什么是自己的长处与优势，什么是自己的不足与劣势。对于任何一个不愿在竞争中失败的人来说，这无疑是走向成功的一条重要途

径。它既可以使一个人在同他人竞争中扬长避短，也能使那些希望自己更上一层楼的人把时间和精力投入到最需要的方面，收到事半功倍的效果。所以，考试也是帮助人们在竞争中立于不败之地的必要条件之一。

（三）考试是人事管理现代化、信息化的重要组成部分

考试一种关于被考者有关情况的信息。这种信息对于人事管理部门和人事管理工作来说，则是科学地实施人事管理的依据。在现代社会系统的人事管理部门中，考试不再仅仅是一种手段或工具，它已成为整个人事管理工作的一个重要组成部分。其具体标志是，在国家机关、省（市、自治区）一级的人事管理部门中，目前已有常设的专司人事考试工作的机构。就是在高等院校、科研院、以及大型企事业中，也有常设的或临时性的专司人事考试工作的机构。此外，社会上已出现以研究考试（包括人事管理考试）为职业的专门机构和科技人员。另一个标志是"人事档案"和"人才信息库"的出现。"人事档案"虽然出现得比较早，但真正成为人事管理部门必不可少的且真正可以引为依据的"人事档案"，则是近现代的事情。至于"人才信息库"的出现，即使在发达的工业化国家，也不过一、二十年的历史。

如果说社会系统中的人事部门设立专司人事考试的机构和出现"人事档案"、"人才信息库"，只是在形式上使考试成为人事管理现代化的组成部分，那么，人事管理工作中的任何决策，都只能来源于反映有关人员情况的信息，而考试正是提供这种信息的根本途径，从这个角度讲，考试便在实质上已成为现代社会系统人事管理工作的重要组成部分。

现代人事管理，又是一个不断地收集信息、分析研究信息并作出决策的过程。这个过程一般分为三个阶段：信息收集阶段、信息分析综合阶段和决策实施阶段。人事管理考试属于第一个阶段，是整个人事管理工作的基础。在人事管理的信息中，各种各

样的人事考试所提供的信息最多、最具有科学性，因而也最重要。在实际的人事管理工作中，通过对这些信息的整理，便成为社会系统中个人档案的另一个重要组成部分——人才档案。这对人员的培养、选拔、晋级和评定技术职称等，都是十分重要的依据。

电子计算机在社会人事管理中的应用，是人事管理技术手段的重大飞跃。随着计算机人事管理系统和更大规模的各种“人才信息库”的建立，人事管理和人才管理的“信息化”要求愈来愈高。为了适应这种新的形势和新的情况，各种各样关于收集人才情况的考试也必然愈来愈经常化、制度化。随着这种经常化、制度化进程的推进，人事考试、人才考试的本身也将愈来愈科学化。可以肯定，在未来社会的人事、人才管理中，考试的作用将更为突出，考试的地位还将进一步提高。

三、人事管理促进考试结果“社会价值”的实现

对于人员的管理和人力资源的利用与开发来说，考试结果在未被采用之前，它只具有理论上的价值。要使考试结果的“理论价值”转变为“社会价值”，还需通过人员管理和人力资源利用与开发的社会实践。这种社会实践活动就是人事管理。也就是说，人事管理可以使考试结果的“理论价值”转化为“社会价值”，成为推动社会物质文明和精神文明建设的现实力量。这种价值的转化，主要反映在对人的价值的充分实现，即对人力资源的合理利用和有效开发。人事管理使考试结果由“理论价值”转化为“社会价值”，表现为一个渐进的过程。当考试结果已定的情况下，这个过程的进展，便取决于人事管理部门对考试结果采用的迟早和快慢，而转化深度（亦称“转化率”，系指一个人考试结果的理论价值与其转化为社会价值之比）则决定于人事管理

部门在其决策中对考试结果利用的程度。一般情况是，考试结果的“理论价值”转化为“社会价值”的“转化率”，往往与人事管理决策中利用考试结果的程度成正相关。反过来，当人事管理决策对考试结果全部利用时，其“转化率”就决定于考试结果的可靠性和有效性。

四、考试与人事管理的互补与互促

人事管理的实践，既检验考试的科学性程度，又不断推动考试理论和实践的发展。当人事管理部门在以考试结果为依据安排和使用人员的实践过程中，证实经考试选录的人员在其工作实践中表现出来的知识、能力、思想意识等，与其考试的结果相吻合时，证明考试结果真实地反映了被测人员的实际水平，这种考试就是富有科学性的。如果考试所选人员所测得的个人成绩，与被测者在工作实践中表现出来的知识、能力和思想意识等，只是部分相吻合，则说明考试缺乏科学性，其测试结果不甚准确可靠。如果被测者在其工作实践中表现出来的知识、能力和思想意识等，与他们考试的结果完全相反，这种考试则无科学性可言。所以说，人事管理的实践是检验考试科学性及其结果客观性的唯一标准。

正因为人事管理的实践具有检验考试优劣及其结果可靠性的功能，所以，随着人事管理实践的深入和科学化程度的提高，必然推动和促进考试理论和实践的发展。作为现代人事管理活动重要组成部分的人事考试，对于用人过程中所提出的应由考试解决的问题，也必须研究解决。而在这种不断解决用人实践所提问题的过程中，考试的理论与实践又不断地得到丰富、发展和完善。

人事管理活动的实践还证明，人事考试的理论与实践，既有需要人事管理实践检验的一面，又有指导人事管理实践的一面。从某种程度上讲，人事考试对于人事管理中的用人具有一定的“超前性”。因为唯有这样，才能不断地和越来越科学地为人事

管理中的用人提供客观依据。而随着考试理论的发展和实践的丰富，又为加速人事管理的科学化创设条件。这正是考试与人事管理之间相互依赖、互相促进的辩证关系。

第二节　人事管理系统考试的组织机构及其职能

在现代人事管理系统中，由于考试是准确选人合理用人的基本手段和前提条件，所以，专司人事考试的机构相继建立，专司人事考试工作的人员逐年上增。为了正确认识人事管理系统考试机构的性质与任务，充分发挥人事考试的作用，本节将着重探讨人事管理系统考试机构的设置原则、组织形式、人员结构及其职能。

一、人事管理系统考试机构的性质

一般说来，现代人事管理系统的考试机构具有以下性质：

首先，广义上的人事管理考试机构，由对人事考试的管理和人事考试两个部分组成。各级政府人事部门的人事考试机构，虽然它的主要任务和职能，是代表一级政府对人事管理工作中的考试进行组织和管理，如制订有关的方针、政策、条例等，似乎在一定程度上具有权力机构的性质，但由于它更多的是从事考试业务的管理，对人事考试的方针、政策等并无立法权，故属管理机构的范畴。企事业单位的人事考试机构，也有一个根据政府人事考试的方针、政策等，确定本系统具体实施办法与措施的问题，但由于其主要职能是组织并主持本系统的考试，它也是一种具有管理性质的考试业务机构。总之，不管是政府人事部门的考试机构，还是企事业单位的人事考试机构，都不是权力机构，而是具有某种行政管理职能的业务机构。

其次，人事考试机构对具体的人事决策无决策权。但在人事管理的决策者作出有关人事问题的决策之前，亦即决策过程中，可以而且必须进行参谋。所谓参谋，就是提供本机构所掌握的有关人员的情况，诸如政治主张、思想品质、道德修养、专业特长、工作能力、身体素质，等等；在对有关人员的情况进行了全面而客观的综合分析、比较研究之后，提出可供决策者参考的建议；当决策者已作出决策之后，如认为某一决策有不妥之处，可根据掌握的情况提请决策者考虑。

二、建立人事管理考试机构的原则

人事考试是一项科学性、政策性很强的工作，所以在建立人事考试机构时，应坚持以下三条基本原则：

（一）与本系统所管理的人员范围保持一致

这一原则包括两个方面的内容：一是考试机构的规模与所管理人员数量上的一致性。如果考试机构规模太大，而所管理的人员人数不多，便会人浮于事，影响工作效率。反之，如果所管理的人员人数很大，而考试机构的规模很小，则又可能因力所难及而顾此失彼，影响工作效益。因此，人事考试机构的规模应视所管理人员的人数确定，两者之间保持一定的比例，宜大则大，宜小则小。至于两者之间多大比例为宜，尚待进一步研究和实践。二是考试机构人员的专业结构与所管理人员专业结构应有一致性。本系统有多少种专业人员，考试机构中就应配备相应的专业人员，至少系统中一些主要的专业或科类应有相应的考试人员。如果让一个对施考专业一窍不通的人员去主持该专业人员的考试，无论此人工作能力多大，组织能力多强，也不大可能比由本专业人员主持、组织这种考试的效果好。

（二）与本系统人事管理所要进行考试的范围保持一致

现行人事管理考试涉及范围宽广，包括人的德、学、才、识

体的各个方面。况且，不同的社会系统人员的专业构成不同，对人员全面检测的要求和侧重面也不完全一样。所以，在建立人事管理考试机构时，应充分考虑与本系统人事管理所进行考试范围的一致。唯有如此，才有利于提高考试工作的效率。

（三）维护考试机构的权威性

一般情况下，考试内容、方法的科学性和考试结果的可靠性，是与考试机构本身的权威性密切关联的。而考试机构的权威性只能由两个方面决定：一是本系统的决策者赋予它的权威性，二是考试机构所施考试的质量具有的权威性。在这两个方面，决策者所赋予的权威性往往只具有行政上的意义，考试机构所施考试本身的质量起决定性作用，任何考试机构的权威性，无不取决于考试质量的可靠和有效。因此，人事考试机构的建立，虽应赋予考试机构行政上的权威性，但关键是要选择、配备好考试机构的工作人员，他们必须有胜任考试业务、确保考试质量的实际能力。否则，考试机构将如同虚设，即使本系统的决策者赋予它行政上的权威性，也会因其考试质量的低劣而丧失。

三、人事管理系统考试机构的类别与职能

（一）人事管理系统考试机构的类别

根据各社会系统人事管理的实际情况和需要，考试机构的组织形式在不违背上述原则的前提下，可因地、因事制宜，灵活多样。大体上可分为以下几种类型：

1．常设性考试机构。

这是一种固定的考试机构，适用于规模较大的社会系统。例如，各级政府的人事管理部门和大型企事业人事管理部门，可设人事考试委员会、考试局、考试办公室、考试科或考试组。这类考试机构的特点是，工作具有连续性、稳定性，既从事考试的组织、实施工作，也开展一些关于考试理论方面的研究探讨工作。

2．临时性考试机构。

此类考试机构，一般是为完成某一种或某一阶段的人事管理考试任务而设立的。这种形式既适用于大的社会系统，也适用于较小的社会系统。其组织形式可以是委员会，也可以是办公室或小组。其特点是灵活性大、实用性强。

3．特设性考试机构。

这类考试机构的特点是它的专门性或专业性。所谓专门性或专业性，是指专司某一种专业或某一种特长、技能的考试。这类机构根据实际情况和需要而设，可以是固定的，也可以是临时的。由于它的主要目的是为了发现、选拔一些具有特殊才能的专门人才，属于一种特设的考试机构。

4．综合性考试机构。

这类考试机构原则上只适用于国家各级政府人事部门。其特点是有利于对各级各类人员进行全面系统地考核，信息反馈快，能大面积地发现不同类型的人才。综合性考试机构可以常设，也可以是临时性的。

以上四类考试机构常有这种情况：除机构本身的正式工作人员外，还不时聘请或借用一些有关人员参加一段时间的工作。是否有必要把出现这种情况的考试机构另成一类，例如称之为弹性考试机构呢？从其组织活动形式的角度上讲，另成一类也未尝不可。但考虑到这种现象在常设性、临时性、特设性和综合性考试机构中都可能出现，且往往只是一种临时性措施，故无单独归类的必要。

（二）人事管理系统考试机构的职能与任务

人事管理考试机构的职能是：根据国家有关考试法规和人事管理的要求，制订和执行本系统人事管理考试方面的条例、方案和措施等，并根据变化了的情况或新提出的要求，负责组织、主持有关考试条例、方案和措施的修改；负责本系统人事管理方面

各种考试的组织与实施；为本系统的人事管理决策提供客观依据和提出有参考价值的建议。

根据人事管理考试机构的性质和职能，人事管理考试机构有以下七个方面的任务：

（1）贯彻执行党和国家关于人事管理方面的方针、政策，并以这些方针、政策为依据，结合本系统的具体情况，起草、制订和修改本系统人事管理考试的有关条例、方案和措施等，经本系统的权力机构或领导批准后，负责组织实施；

（2）根据本系统的人员情况和人事管理的需要，研究并确定人事管理考试的内容、组织形式、实施方案等；

（3）负责具体组织并主持本系统人事管理方面的各种考试；

（4）负责建立并完善本系统人事管理工作所必须的人员考试档案资料；

（5）总结本系统人事管理考试工作中的经验和教训，不断提高人事管理考试的科学性。

（6）常设性考试机构，应坚持对本系统人员、特别是那些对本系统工作成败具有决定性或关键性作用的人员，进行连续性、经常性的考核、考察工作。还应结合工作实际开展一些有关人事管理考试问题的研究。

（7）负责宣传党和国家有关人事管理考试方面的方针、政策，以及本系统人事管理考试工作中的经验和好人好事。

为了有效履行职责，应赋予人事管理考试机构以下权利：

（1）在实施人事管理考试的过程中，有权要求被试者回答主试的质问；

（2）在实施人事管理考试的过程中，有权排除与考试目的无关的任何外来干扰；

（3）若发现考试结果不真实，有权要求对原测试对象重新

进行考试；

（4）在实施人事管理考试的过程中，对违犯考试法规者，根据情节轻重，主考人员有权提出批评或取消考试资格，要求有关组织作出处理，直至提出讼诉。

四、人事管理考试机构的人员结构及素质

（一）人事管理考试机构的人员结构

根据人事管理考试机构工作任务的性质与要求，其人员结构应恰当处理以下三个方面的问题：

1．“软科学”人员和“硬科学”人员要保持一定的比例。

从人事管理考试机构的职能和任务可以看出，这种机构是“软科学”和“硬科学”的综合体。所以，其机构成员的结构也应是“软科学”人员和“硬科学”人员的联合体，并应使这两种人员保持合理的比例。

2．机构成员整体的知识结构，应具有较大的覆盖面。

从人事管理考试机构的知识结构要求看，应主要考虑三个方面：一是本系统人员的主要专业分布情况；二是本系统人事管理考试除主要专业（指本系统）知识以外的内容，三是人事考试管理工作所需要的知识。人事管理考试机构成员合理的整体知识结构，应在这三个方面具有较大的覆盖面，并且覆盖面愈大愈好。

3．机构成员应具有较高的政策、业务水平，组织性、原则性强。

由于人事管理考试是一项政策性很强的工作，任何不负责任的态度、自由主义和自以为是的行为，都可能造成人员使用、人才选拔上的失误，甚至影响到安定团结。所以，人事管理考试机构的成员，必须具有较高的政策水平、严密的组织纪律性、坚定的原则性和高度的责任感。那种工作马虎、对一切都无所谓的人，是不适宜在人事管理考试机构中工作的。

（二）人事管理考试机构工作人员的素质要求

人事管理考试机构的工作人员，应具有以下四个方面的素质：

1．政治素质：有较高的政策水平、很强的组织纪律性、坚定的原则性和高度的工作责任感；

2．文化素质：懂得人事管理学、人才学、心理学、行为科学和考试学方面的基本理论与知识，并能将所学知识创造性地用于工作实际；

3．业务素质：熟知有关考试的方针、政策和法规，熟悉考试设计、实施的基本程序，具有组织和实施考试，解决考试过程中的有关方法、技术问题和处理突然事件等方面的能力；

4．热爱人事管理考试工作。

第三节 人事管理系统考试的范畴、类别、原则与方法

人类社会的考试，可分为社会考试和学校考试两大系统。人事管理考试是社会考试系统的主体，其范围之大、类别之多、影响之广，远非其它各类考试可比。同时，它作为人类考试群体中一个重要分支，有其特定的性质和任务。因此，人事管理系统考试的设计与实施，有其固有的原则和方法。

一、人事管理系统考试的范畴与类别

正确区分和确定人事管理系统考试的范畴与类别，是考试学的基本任务之一，也是建立和完善考试学理论体系的必需，同时，对实现人事管理的科学化、现代化，将产生积极的促进作用。

（一）人事管理系统考试的范畴

现行考试不但种类多、范围广，而且几乎涉及人群的每一个层次——少年、青年、中年和老人。如何才能正确地区分和确定人事管理系统考试的范畴呢？我们认为，刘分考试范畴的根本依据是考试的目的和对象。人事管理系统考试的目的，或是评量在职人员的技术、业务水平、政治思想水平、身体素质，以加强或改进管理；或是甄选新的工作人员，以充实或更新职工队伍；或是选拔某种特殊人才，以解决某部门工作急需，等等。人事管理系统考试的对象，是在职人员、劳动后备军和其他具有某种工作能力的人。这就是人事管理系统考试的特殊点，也是人事管理系统考试与其他考试的区别点。所以，凡是以上述对象和目的设计、实施的各种考试，均属人事管理系统考试的范畴。具体包括五大方面的考试：

1．招工、招聘考试。

这是企事业单位为了招收新的工人或招聘新的专门人才，如律师、顾问、教师、工程师、技师、护士、厂长、经理等，所举行的一种人事选拔考试。其特点是对应考者有一定的条件要求，有一定的竞争性。按照规定的条件和要求，对应招、应聘者进行测试，然后根据应试者的测试成绩，择优录用。

2．为了解、掌握在职人员基本情况的考试。

这是现代人事管理系统中形式最多、范围最广的一种考试。其目的是了解、掌握在职人员的基本情况，以利于加强管理。考试内容包括：政治思想、工作成绩、业务能力、专业技术、身体素质等，以及人事管理部门认为应该了解或掌握的其他内容。

对于上述考试内容，可以是全部考，也可以只考其中一部分。其考试形式可以是书面的，也可以是口头的，可以是总结式的，也可以是问答式的，可以独立进行，也可以相互评论的形式进行；可以公开施考，也可以不公开测试等等。

3．评定技术职称或晋升职务的资格考试。

这是在各种专业技术人员中进行的一种人事管理考试。其目的是对每一个参加考试的专业技术人员的工作成就、学术水平、工作能力等，作出公正客观的评价，并以此为依据确定其相应的职称或职务，便于在工作中正确使用这些专业技术人员，充分发挥他们的聪明才智和技术专长。这种考试的内容一般是根据被试所从事的工作或专业要求确定。

4．甄别、选拔各种干部的考试。

这也是现代人事管理系统中常用的一种考试。考试内容往往是根据所要选拔的干部类别及标准而定。如选拔政工干部，考试的内容多侧重政治思想、道德品质、工作方法、政治理论修养等；如选拔业务管理干部，考试内容又侧重于业务管理能力、业务理论修养、政策理论水平和组织能力等；如果选拔领导干部，考试的内容则应以决策能力、组织指挥能力、领导科学理论素养等为主。这种考试的特点是，除特殊情况外，一般不公开进行。

5．甄选某些特殊人员的考试。

此种考试多用于政府部门或企事业单位，目的在于甄别、选拔某些具有特殊要求的人员。例如为了振兴我国体育事业，国家专业运动队甄别、选拔专业运动员的考试，就属于这种考试。这种考试的特点是考试内容特定，并且难度往往较大，对应试者的条件要求较严。

最后顺便指出，现代社会中政府部门或企事业单位进行的人员测评活动，也是现代人事管理考试的一种形式。这是因为，第一，测评对象主要是各行各业的在职人员；第二，测评的目的，是对各类人员的德、学、才、识、体进行定性和定量相结合的测量与评定，作为选拔、任用、培养和人员组合的依据，以求取人员的能力特点与工作要求相匹配，进而实现人尽其才、才尽其用以及人员管理合理化的目的。这一切，均与我们对人事管理系统

考试的定义相一致。所以，人员测评活动当归入现代人事管理考试的范畴。

（二）人事管理系统考试的类别

人事管理系统考试，根据考试的性质和目的可分为两大类：评价考试和选拔考试。

1．评价考试。

评价考试的主要目的，旨在对有关人员的某一个方面、几个方面或所有方面的情况作出评价。例如科技专业人员评定职称的资格考试，青年徒工学徒工的定级考试，人事管理部门为了解、掌握本系统工作人员某些情况而进行的考试等，都属于评价考试。这类考试一般具有以下特点：

（1）被试者不存在淘汰的问题；

（2）适用范围广：适用于工人、科技人员、政工人员、领导干部等各类人员；

（3）形式灵活多样，书面、口头、操演、公开、隐蔽、个体、团体、定期、不定期等形式，都可采用；

（4）内容不受限制：可以是综合性的全面考试，也可以是某一个方面的专题考试。

评价考试，是人事管理部门建立和完善人员档案的重要手段和方法。正确运用评价考试，对于人事管理部门知人、育人、用人，充分开发人力资源，提高人才使用的经济效益，都有不可低估的作用。

2．选拔考试。

属于人事管理选拔考试的有招工考试、招聘考试、甄选某些特殊人才的考试和选拔干部的考试等。其基本特点是：

（1）考试结果直接导致被试者生活环境或工作岗位的变化；

（2）适用范围有一定的局限性，如群众对领导、下级对上

级就不适宜采用这类考试；

（3）某些特殊性质的选拔考试不能在公开的情况下进行，如选拔领导干部的考试，一般不公开进行；

（4）考试内容的专业性较强；

（5）除个别情况（如数量大的招工）外，考试的规模较小。

在人事管理部门中，选拔考试除用于招收新工人外，还可以广泛运用于对各种专门人才的甄别和选拔。只要考试得体、方法恰当，一般都可以收到较好的效果。

3．人事管理系统评价考试与选拔考试的关系和异同。

人事管理的评价考试与选拔考试之间的关系十分密切。在实际的人事管理中，有时往往是评价与选拔同时处于一个过程之中。还有，即使是评价考试，在客观上有时也可能产生选拔考试的结果，即导致有关人员生活环境或工作地位的变化。

一般说来，选拔考试中包含有评价的成份，而且任何选拔考试最终结果的产生，都要以评价为基础。所以从某种意义上讲，评价考试是选拔考试的基础。例如某大学需要甄选一名党委宣传部长。条件已定。学校组织部门经过各方面的甄别，从全校的政工干部中确定了3个候选者（这已是评价的结果）。然后按照已经确定的条件和要求，再运用多种形式对这三个候选者进行考试，最后选出其中最符合标准的一个任宣传部长。这显然是一种选拔考试，但选拔的过程又包括一系列的评价。

反过来，评价考试是否也具有某些选拔考试的功能呢？从目的和动机上讲，评价考试不会直接产生选拔考试的结果，因而也不会有选拔考试的功能。但各种事物之间不可能是绝然孤立的，评价考试随着数量的变化，往往会出现质的飞跃，也会导致选拔考试的结果。因为随着一次又一次的评价考试，人事部门对被评价者的印象不断深化，如被评价者中确有才华者或德才兼备者，

而本系统又正需要那样的人才，这位人才就可能不经过专门的选拔考试而被安排到相应的职位。

人事管理中的评价考试和选拔考试之间虽然有着密切的联系，但两者之间也有明显的不同之处，这是两种考试既不能等同，也不能互相替代的原因。评价考试与选拔考试的不同之处在于：

（1）目的不同。评价考试的目的是对被试者的实际情况作出评价，而选拔考试的目的是从被试中选优。

（2）产生的效果不同。评价考试对被试者来说不存在淘汰的问题，也不直接改变被试者的生活环境或地位。而选拔考试则恰恰相反，被试者中既有被淘汰者，同时必然改变入选者的生活环境或地位。

（3）适用的范围不同。评价考试适合各种对象，既适用于一般人员，也适用于领导干部，包括下级对上级。而选拔考试就不适用于下级对上级或群众对领导。

综上所述，评价考试和选拔考试是人事管理系统两种基本的考试类别，它们之间既有密切联系的一面，又有明显的不同之处。在人事管理考试的实践中，既不应将两者截然分开，也不应将两者相互替代。正确地运用评价考试和选拔考试，并使二者有机结合，互为补充，加上科学的考试内容、正确的考试方法和恰当的考试形式，定可产生积极的效果。

二、人事管理系统考试的原则与方法

大力开发潜在的人力资源，充分实现人才的价值，是国家兴盛、社会进步的基本保证。正确地甄选和使用人才，则是实现此目的的重要前提。然而，要做到人才甄选准确、任使得法、效益良好，就必须有正确的考试原则和科学的考试方法。

（一）甄选、使用人才的基本原则

下述诸点，是甄选、使用人才必须遵循的基本原则：

1．坚持德、学、才、识、体全面检测，但不"求全责备"。

"金无足赤，人无完人。""求全责备"同样为甄选、使用人才不可取。对于人才的甄别和选用，既要坚持全面考察，又要防止因瑕疵而失白璧，应着重于本质、主流。譬如某一位科学家，过去曾说过一些错话，或者在生活小节上存在一些问题，但他对社会主义祖国始终是热爱的，而且在学术上确有很高的造诣，事业进取心强，同时具有很强的组织管理能力。对于这位科学家，我们就应大胆起用，将其安置到能充分发挥他的聪明才智的岗位上，而不应闲置，致使他不能尽其才能于社会。

2．坚持"唯贤是用"而不论亲疏。

唐太宗李世民说过："吾为官择人，唯才是与。苟或不才，虽亲不用；如其有才，虽仇不弃。"李世民的这种人才思想和用人原则，直到今天，仍不失其光辉。今天，在我们为现代化宏伟事业甄选和使用人才时，更应坚持"唯贤是用"而不论亲疏。我们的组织部门、人事部门的同志，肩负着为党和国家甄选和使用人才的重任，为了党的事业，为了国家民族的兴盛，在甄选、使用人才时，应该"内举不避亲，外举不避仇"。

3．坚持能级原理，力求动态对应。

所谓能级，是指人们开展工作的能力或能量大小的等级。显然，某一个人在某一个时期的能力是一定的，而不可能是任意的。这种按照人们实际工作能力的大小，将其分成若干等级，并以此为依据恰当分配工作的方法，现代管理学上称之为"能级原理"。这是人才甄选和使用理应坚持的基本原则。

坚持能级原理，力求动态对应包括两层内容。

第一，在甄选、使用人才时应通过各种手段、途径和方法，对被甄选、使用者的实际能力进行考察，并作出符合本人实际情

况的评价，以准确确定其能级。然后，视其能级的高低，安排与之相适应的工作职位。这是一项极为严肃而又科学性很强的工作，在执行的过程中，决不能带有任何感情色彩和私心杂念。

第二，一个人的一生，其能级也不是一成不变的。在正常情况下，一个人能级的变化一般呈"正态"型。即：当一个人刚刚开始工作时，由于缺乏经验、社会阅历浅，其能级一般较低。随着年龄的增长，工作经验的积累，社会阅历的丰富和其他知识的增多，其能级将逐步上升（少数人会出现能级"跳跃"现象），当一个人的年龄、精力、经验、知识和思维能力等，均达到并处于最佳状态时，其能级则将跃迁到能力所能达到的极限，即此人能级之最高阶段。过了这个阶段，当年龄达到某一较大值时（一般为55岁左右），由于生理上的原因，其精力旺盛度、思维敏捷度和反映灵敏度等渐趋下降。虽然此时的经验和社会阅历比以前更丰富，但其能力的等级则不可避免地会降低。年龄越大，能级也就愈低。因此，对一个人能级的甄别，切不可一劳永逸，而应根据人的能级变化规律，分阶段进行甄别、确定。并以此为依据，在其不同的阶段，视其当时的能级决定其工作职位。干部终身制之所以必须废除，这也是重要的原因和理论根据之一。

（二）甄别、选拔人才的基本方法

"千兵易得，一将难求"。这既说明了人才之宝贵，也说明了发现人才之不易。上面我们讨论和研究了甄别、选拔人才应该遵循的基本原则，这对于发现人才、使用人才和管理人才无疑是十分重要的。但这些原则所解决的只是甄别、选拔人才的指导思想，而任何人才的发现与获得，都离不开具体的措施和方法。

甄别、选拔人才的方法，是人才学研究的重要内容，更是考试学研究的基本内容之一。因而在考试学中占有重要的地位。这主要表现在两个方面：其一，关系到甄别、选拔人才的基本原则（即指导思想）的贯彻落实。甄别、选拔人才的基本原则和甄

别、选拔人才的基本方法，是人才管理中一个问题的两个方面或两个层次，两者互相关联，相互影响。甄别、选拔人才的基本原则是甄别、选拔人才的理论问题，其任务是解决甄别、选拔人才活动的指导思想，为甄别、选拔人才的活动确定必须遵循的准则，指导甄别、选拔人才方法的选择与具体实施。而甄别、选拔人才的方法，则是甄别、选拔人才的具体措施或途径，表现为具体行动的一种过程。这种行动过程，受甄别、选拔人才基本原则的制约。所以，它总是正确或不正确地反映着甄别、选拔人才的指导思想。其二，甄别、选拔人才的方法，直接关系到人才本身的前途和所选人才的质量。因为甄别、选拔人才活动的结果，无一不是对被甄别、选拔者作出评价，而被甄别、选拔者是否是人才，是一般人才还是杰出的人才，又往往是以甄别、选拔活动的结果为依据的。如果结果的客观性程度高，由此作出的结论便切合实际，如果结果的客观性程度较低，由此作出的结论就可信度低。无论是前一种情况还是后一种情况，都直接关系到人才本身的前途和所选人才的质量，对国家和民族的事业有着直接影响。这就要求甄别、选拔人才的方法必须具有高度的科学性、严肃性，以保证真正的人才不被埋没，平庸之辈无法鱼目混珠，在位者无庸才，山野处无遗士。

纵观人类甄别、选拔人才的方法，虽代有翻新，古今有别，中外有异，但归结起来，其基本方法不外乎三大类：

第一类，直接考试法。

用直接考试的方法甄别、选拔人才，开始于奴隶社会，如斯巴达农业贵族的武士考试，雅典商业贵族的学校考试，我国西周时期的选士考试等。到公元7世纪初，直接考试法在我国更为完善，如科举考试，就是我国从隋唐至清末历代王朝甄选人才的基本途径和方法，并为东西方国家相继效仿。时至近现代，直接考试法已成为人类甄别、选拔人才的主要方法。

运用直接考试法甄别、选拔人才的最大特点是，在特定的条件下，甄别、选拔人才的标准统一，除个别特殊情况外，一般可以做到使被甄别、选拔者在分数面前人人平等。但也有明显的不足之处，仅仅靠一、两次考试的成绩作为选用与否的决策依据，有时并不能真正反映一个人的实际才能，甄选人才之偶然性实难避免。

第二类，考察荐举法。

这是人类社会最早用于甄别、选拔人才的方法，其历史比直接考试法更为悠久，可追溯到原始社会的后期。我国历史上的荐举制、内正法、乡举里选制、九品中正制、连坐保举制、制举制，以及我国历代统治者对各级官吏的考绩、考课制度，如上计、刺察制、司考、校考制、考满、考察制、察计制等，都属于这类方法。在现实社会，考察荐举法仍是甄别干部、选拔人才的基本方法之一。

运用考察荐举的方法甄别、选拔人才，由于要对甄别、选拔对象进行较长时间的、多方面的观察和分析，所以在正常的情况下，往往能够对被甄别、选拔者作出比较客观的评价。其不足之处在于：甄别、考察的结论的真实性，往往受到考察者、推荐者好恶、思想作风等因素的影响。特别是在社会风气不好或社会统治集团作风腐败的时候，考察荐举人才的方法，就会成为官吏们结党营私、推亲荐故的借助手段。

第三类，综合法。

所谓综合法，就是将直接考试法和考察荐举法同时运用于人才的甄别或选拔。

由上述可知，直接考试法的长处正好是考察荐举法的不足之处，而考察荐举法的优点又恰恰是直接考试法的缺点。可见，这两种甄别、选拔人才的方法具有互补的性质。综合法正好是兼两家之长而又避两家之短，在现行甄别、选拔人才的各种方法中，

它是一种科学性、客观性和可靠性都比较高的方法。随着综合法的不断改进和完善，它必将成为人们甄别、选拔各类人才的主要方法。

第四节　人事管理系统考试的基本模式

建立人事管理系统考试的基本模式，是人事管理考试的基础。它的作用和功能有如工业设计中的规范，数学中的公式。同时，建立人事管理系统考试的基本模式，是将电子计算机应用于人事管理考试的重要条件，因而也是使人事管理考试方法和手段现代化的重要前提。本节将着重讨论人事管理系统考试建模的要求，以及人事管理考试的有关注意事项，并以表格形式介绍十类人员之人事管理考试的模式。

一、人事管理考试建模的基本要求

所谓人事管理考试的模式，就是用若干具体指标（或参数）组成一个能反映一般人员（指具有工作能力且在职的劳动者）或一类人员的基本情况的目标参照系，用于对一般人员或一类人员的考试。一个好的人事管理考试模式，应具备三性：

（一）准确性

所谓准确性，包括两层意思；一是构成模式的每一个指标都能准确地反映人员的某一个方面或某一种基本情况；二是整个模式应能准确地反映被试者的各主要方面的基本情况，或某种模式所希望说明的那些内容。为达此目的，对构成模式的每一个指标都应进行认真的、反复地斟酌，并给它们以明确的目标和确定的内涵。第二，每一个指标必须具有相对独立性，不允许有互相覆盖或可互相取代的现象。第三，由一个个相对独立的指标组成的模式，必须是一个有机的整体。

（二）简明性

一般说来，人事管理考试模式的指标越多、越具体，其可靠性就越高。但一个好的模式要求具有一定的简明性。所谓简明性，一是组成模式的指标要简单明确，使人一目了然；二是整个模式要简单明晰，条目层次清楚。显然，构成模式的指标不宜太多。否则，模式系统就难免庞杂，而且各指标之间，也很难避免互相覆盖的现象，使其独立性受到影响。

（三）可行性

建立模式的目的是要用于实际，减少以至避免实际工作中的一些重复性劳动，并使其结果具有更高的可信度，从而达到既提高工作效率又提高人事考试科学性的效果。所以，一个好的模式必须紧密结合实际，具有可行性。如果在实际使用中行不通，即使最完美的模式，也是毫无意义的。

人事管理考试模式的可行性，主要是指构成模式体系的指标参数的可采度。模式中的指标参数的可采度愈高，模式的可行性就愈好。人们在建立具体的人事管理考试模式时，总是追求高度准确、简明、可行的理想目标，但要这三个方面都达到“最佳值”，事实上有困难。因为人事管理考试模式的准确性、简明性和可行性是互相制约的。我们在对它们作出决策时，不要盲目追求“最佳值”，能取“满意值”即可。

二、十类人事管理考试模式

根据有关的调查和研究，我们试建了十类人员人事管理考试的基本模式，既作为建立人事管理考试模式的举例，也供本书有关读者作参考。

（一）关于建立人事管理考试模式的几点说明

首先，在现代社会中，既具有劳动能力又在职的人员情况是千差万别的，这里虽然主要是指其工作岗位方面的不同，但也包

括人员自身方面的差异。所以，不可能建立一个适用于社会一切人员的万能人事管理考试模式，而只能建立以某一工作岗位人员为基准的或以某一类型人员为基准的具体模式。

其次，在建立某一种人事管理考试模式之前，必须事先进行深入、细致、认真、严肃的调查研究，弄清该模式所必须反映的某一工作岗位人员或某一类人员在人事管理方面的内容，以便准确选择和确定构成模式的各种指标。

第三，在建模过程中，除应充分注意前面所论述的三条基本原则外，还应充分考虑每一个指标的测试形式和评定结果的方法，以便在实际运用模式时，能按照统一的规范进行，不致因盲目而影响考试效果。

（二）十类人员人事管理考试的基本模式

为进一步说明人事管理考试建模的要求，以及各类考试模式的基本内容，并为人事管理考试模式的建立提供不同类型的参考样本，我们特意试建了科研、管理、政工、领导干部、教师、司法、检察、外贸、军事、文学美术、电子计算机等十类人员人事管理考试的基本模式，现分列于后。

表11-1　　科学研究人员人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．事业心	评定法	定性法	4．专业知识	答卷法	定量法	8．创造能力	评定法	定性法	12．科研成果	评定法	定性法
	2．进取性	评定法	定性法	5．基础理论知识	答卷法	定量法	9．选题能力	评定法	定性法			
	3．科研道德	评定法	定性法	6．外语知识	答卷法	定量法	10．表达能力	评定法	定性法			
				7．情报检索知识	答卷法或问卷法	定量法	11．思维能力	评定法	定性法			
选考部分	1．协作性	评定法	定性法	3．知识更新	评定法	定性法	4．容纳信息能力	评定法	定性法			
	2．坚韧性	评定法	定性法				5．学术鉴别能力	问卷法或演示法	定性法			

表11-2　管理人员人事管理考试的基本模式*

人员类别		素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
技术管理人员	必考部分	1．政策观念 2．事业心 3．责任感	评定法 评定法 评定法	定性法 定性法 定性法	4．专业知识 5．知识面	答卷法 答卷法或问卷法	定量法 定性法	6．科技方面的鉴别能力 7．组织协调能力	评定法 评定法	定性法 定性法	8．工作效率 9．工作效果	评定法 评定法	定性法 定性法
	选考部分	1．技术、经济观念	评定法	定性法	2．思维力	答辩法或交谈法	定性法	3．灵活性 4．信息沟通能力	评定法 评定法	定性法 定性法	5．科技成果及社会、经济效益	评定法	定性法
行政管理人员	必考部分	1．政策观念 2．法制观念 3．群众观念 4．责任心 5．纪律性	评定法 评定法 评定法 评定法 评定法	定性法 定性法 定性法 定性法 定性法	6．工作经验 7．现代科学知识	评定法 答卷法	定性法 定量法	8．处事能力 9．表达能力 10．信息沟通能力	评定法 评定法 评定法	定性法 定性法 定性法	11．工作效率 12、工作效果	评定法 评定法	定性法 定性法
	选考部分	1．公道	评定法	定性法				2．综合分析能力 3．辅助决策能力 4．控制能力	评定法 评定法 评定法	定性法 定性法 定性法	5．社会效果	评定法	定性法

*这里的管理人员不含领导干部

表11-3-1　　政工人员（宣传）人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．政治坚定性	评定法	定性法	6．马列主义理论知识	答卷法	定量法	8．观察能力	评定法	定性法	11．工作效果	评定法	定性法
	2．政策水平	评定法	定性法	7．政工宣传知识	答卷法	定量法	9．表达能力	评定法	定性法			
	3．事业心	评定法	定性法				10．说服能力	评定法	定性法			
	4．原则性	评定法	定性法									
	5．理论联系实际的精神	评定法	定性法									
选考部分	1．正直性	评定法	定性法	3．知识面	答卷法或问卷法	定性法	4．思维能力	评定法	定性法			
	2．自我批评精神	评定法	定性法				5．创新能力	评定法	定性法			
							6．沟通能力	评定法	定性法			

表11-3-2 政工人员（组织类）人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．政策性	评定法	定性法	6、组织工作专业知识	答卷法	定量法	8．识别人才的能力	评定法	定性法	12、工作效果	评定法	定性法
	2．政治坚定性	评定法	定性法	7、马列主义理论知识	答卷法	定性法	9．用人之长的能力	评定法	定性法			
	3．事业心	评定法	定性法				10．辅助决策的能力	评定法	定性法			
	4．责任性	评定法	定性法				11．综合分析的能力	评定法	定性法			
	5．原则性	评定法	定性法									
选考部分	1．正直性	评定法	定性法	3．知识广度	答卷法或问卷法	定性法	4．与人相处的能力	评定法	定性法			
	2．平等性	评定法	定性法				5．说服能力	评定法	定性法			
							6．协调能力	评定法	定性法			

表11-4-1　　领导干部（高校校长）人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．政策水平	评定法	定性法	4．领导科学知识	答卷法	定量法	7．决策能力	评定法	定性法	13．学校学术地位	评定法	定性法
	2．事业心	评定法	定性法	5．管理专业知识	答卷法	定量法	8．借织能力	评定法	定性法	14．学校社会影响	评定法	定性法
	3．责任心	评定法	定性法	6．教育科学知识	答卷法	定量法	9．综合分析和解决问题的能力	评定法	定性法	15．个人的社会威信	评定法	定性法
							10．知人和用人能力	评定法	定性法			
							11．学科发展和建设能力	评定法	定性法			
							12．人际关系能力	评定法	定性法			
选考部分	1．为人师表	评定法	定性法	3．知识面	答卷法或问卷法	定性法	4．知识更新能力	评定法	定性法			
	2．自知之明	评定法	定性法				5．演讲能力	评定法	定性法			
							6．协调能力	评定法	定性法			

表11-4-2 **领导干部（厂长、经理）人事管理考试的基本模式**

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．政策性	评定法	定性法	5．领导科学知识	答卷法	定量法	7．决策能力	评定法	定性法	12．经济效益	评定法	定性法
	2．事业心	评定法	定性法	6．经营管理知识	答卷法	定量法	8．识人用人能力	评定法	定性法	13．企业发展	评定法	定性法
	3．进取精神	评定法	定性法				9．组织指挥能力	评定法	定性法			
	4．责任性	评定法	定性法				10．应变能力	评定法	定性法			
							11．开拓创造能力	评定法	定性法			
选考部分	1．自知之明	评定法	定性法	2．知识面	答卷法或问卷法	定性法	4．社交能力	评定法	定性法	6．社会影响	评定法	定性法
				3．人际关系学知识	答卷法或问卷法	定量法	5．协调能力	评定性	定性法			

表11-5-1

教师（大学）人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1. 事业心	评定法	定性法	4. 基础理论知识	答卷法	定量法	9. 演讲能力	评定法	定性法	13. 教学效果	评定法	定性法
	2. 为人师表	评定法	定性法	5. 专业知识	答卷法	定量法	10. 知识更新能力	评定法	定性法	14. 科研成果	评定法	定性法
	3. 责任心	评定法	定性法	6. 教学法	答卷法	定量法	11. 科学研究能力	评定法	定性法			
				7. 大学生心理学	答卷法	定量法	12. 思维能力	评定法	定性法			
				8. 外语知识	答卷法	定量法						
选考部分	1. 政治倾向	评定法	定性法	2. 情报学知识	答卷或问卷法	定量法	4. 思想工作能力	评定法	定性法	5. 师生关系	评定法	定性法
				3. 高等教育学知识	答卷或问卷法	定量法						

表11-5-2　教师（小学）人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．职业事业心	评定法	定性法	5．教学法	答卷法	定量法	8．演讲能力	评定法	定性法	11．教学效果	评定法	定性法
	2．为人师表	评定法	定性法	6．基础知识和基本技能	答卷法和评定法	定量、定性法	9．教导能力	评定法	定性法	12．教育效果	评定法	定性法
	3．责任心	评定法	定性法	7．儿童心理学	答卷法	定量法	10．发现和解决问题的能力	评定法	定性法			
	4．献身精神	评定法	定性法									
选考部分	1．政治倾向	评定法	定性法	2．教育学知识	答卷法	定量法	3．自学能力	评定法	评定法	5．师生关系	评定法	定性法
							4．组织管理能力	评定法	评定法			

表11-6　　司法人员人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．政策水平	评定法	定性法	7．法律专业知识	答卷法	定量法	9．判断能力	评定法	定性法	13．量刑正确率	评定法	定性法
	2．法律观念	评定法	定性法	8．犯罪心理学知识	答卷法	定量法	10．表达能力	评定法	定性法	14．工作效率	评定法	定性法
	3．事业心	评定法	定性法				11．执法能力	评定法	定性法			
	4．责任心	评定法	定性法				12．调查研究的能力	评定法	定性法			
	5．公正性	评定法	定性法									
	6．证据观念	评定法	定性法									
选考部分				1．逻辑学知识	答卷法	定量法						
				2．社会学知识	答卷法	定量法						
				3．知识广度	评定法	定性法						

表11-7 **检察人员人事管理考试的基本模式**

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．政策水平	评定法	定性法	7．法律知识	答卷法	定量法	10．发现问题的能力	评定法	定性法	14．批捕准确率	评定法	定性法
	2．事业心	评定法	定性法	8．检察专业知识	答卷法	定量法	11．判断能力	评定法	定性法	15．工作效率	评定法	定性法
	3．责任心	评定法	定性法	9．犯罪心理学知识	答卷法	定量法	12．表达能力	评定法	定性法			
	4．刚正性	评定法	定性法				13．交涉能力	评定法	定性法			
	5．证据观念	评定法	定性法									
	6．法律观念	评定法	定性法									
选考部分				1．逻辑学知识	答卷法	定量法	4．获得信息的能力	评定法	定性法			
				2．社会学知识	答卷法	定量法						
				3．知识广度	评定法	定性法						

表11-8　　外贸人员人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1. 政策观念	评定法	定性法	4. 外贸专业知识	答卷法	定量法	6. 处事能力	评定法	定性法	11. 经济效益	评定法	定性法
	2. 责任性	评定法	定性法	5. 外语知识	答卷法	定量法	7. 涉外能力	评定法	定性法	12. 工作效率	评定法	定性法
	3. 国格人格	评定法	定性法				8. 思维判断能力	评定法	定性法			
							9. 口头表达能力	评定法	定性法			
							10. 反应能力	评定法	定性法			
选考部分	1. 事业心	评定法	定性法	2. 知识广度	评定法	定性法	3. 灵活性和原则性	评定法	定性法			
							4. 收集信息能力	评定法	定性法			

表11-9-1　　军事人员（参谋人员）人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1. 政治坚定性	评定法	定性法	7. 现代军事科学技术知识	答卷法	定量法	8. 判断能力	评定法	定性法	14. 军事行动的效率和效果	评定法	定性法
	2. 爱国主义	评定法	定性法				9. 反应能力	评定法	定性法			
	3. 责任性	评定法	定性法				10. 辅助决策能力	评定法	定性法			
	4. 牺牲精神	评定法	定性法				11. 信息畅通能力	评定法	定性法			
	5. 事业心	评定法	定性法				12. 军事行动中的创造能力	评定法	定性法			
	6. 组织纪律性	评定法	定性法				13. 知识更新能力	评定法	定性法			
选考部分	1. 国际主义观念	评定法	定性法	4. 知识广度	评定法	定性法	5. 灵活性	评定法	定性法			
	2. 艰苦性	评定法	定性法				6. 动手能力	评定法	定性法			
	3. 沉着程度	评定法	定性法				7. 组织协调能力	评定法	定性法			

表11-9-2　　军事人员（后勤）人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1. 政治坚定性	评定法	定性法	7. 军事后勤专业知识	答卷法	定量法	8. 计划能力	评定法	定性法	12. 工作效果和效率（对战备和军事行动的后勤保障和服务效果）	评定法	定性法
	2. 爱国主义	评定法	定性法				9. 组织协调能力	评定法	定性法			
	3. 责任性	评定法	定性法				10. 辅助决策能力	评定法	定性法			
	4. 事业心	评定法	定性法				11. 处事能力	评定法	定性法			
	5. 服务性	评定法	定性法									
	6. 实干精神	评定法	定性法									
选考部分	1. 国际主义观念	评定法	定性法	4. 现代军事科技知识	答卷法或问卷法	定性法	5. 思维判断能力	评定法	定性法			
	2. 牺牲精神	评定法	定性法				6. 综合分析能力	评定法	定性法			
	3. 组织纪律性	评定法	定性法									

表11-10 文学人员人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．政治热情	评定法	定性法	5．文学专业知识	答卷法	定量法	8．文字表达能力	评定法	定性法	12．成果	评定法	定性法
	2．事业心	评定法	定性法	6．知识广博性	答卷法或问卷法	定性法	9．观察判断能力	评定法	定性法	13．社会效果	评定法	定性法
	3．群众观念	评定法	定性法	7．形象逻辑学	答卷法	定量法	10．构思能力	评定法	定性法			
	4．正直性	评定法	定性法				11．形象思维能力	评定法	定性法			
选考部分	1．自我批评精神	评定法	定性法	4．美学	答卷法	定量法	6．深入生活的能力	评定法	定性法			
	2．道德观念	评定法	定性法	5．社会心理学	答卷法	定量法	7．反应能力	评定法	定性法			
	3．刻苦性	评定法	定性法									

表11—11　　电子计算机软件人员人事管理考试的基本模式

	素质结构	考试方法	评定方法	知识结构	考试方法	评定方法	能力结构	考试方法	评定方法	绩效结构	考试方法	评定方法
必考部分	1．事业心	评定法	定性法	5．系统专业知识	答卷法	定量法	7．计划能力	评定法	定性法	14．工作质量	评定法	定性法
	2．责任性	评定法	定性法	6．项目管理专门知识	答卷法	定量法	8．解决实际问题的能力	评定法	定性法	15．工作效率	评定法	定性法
	3．服务性	评定法	定性法				9．划分系统职能的能力	评定法	定性法	16．效益	评定法	定性法
	4．主动性	评定法	定性法				10．认定和提出目标的能力	评定法	定性法			
							11．发现和分析问题的能力	评定法	定性法			
							12．创新能力	评定法	定性法			
							13．知识更新能力	评定法	定性法			
选考部分	1．政治倾向	评定法	定性法	2．知识广度	答、问卷法	定性法	4．应变能力	评定法	定性法	8．用户反映	评定法	定性法
				3．服务心理学	答卷法	定量法	5．组织协调能力	评定法	定性法			
							6．监督能力	评定法	定性法			
							7．交际能力	评定法	定性法			

三、人事管理考试注意事项

为了保证考试结果的可靠性，在设计、实施人事管理考试时，应注意以下三个问题：

（一）不同劳动的特点

不同类型的劳动，各有其不同的特点，因而在人事管理上对从事不同类型劳动的人员的要求，应互有区别。反映在人事管理考试上，即应以不同劳动的特点为依据确定考试的内容、重点、与之相适应的形式和方法。

现代社会的劳动可分为五大类：物质生产劳动、知识生产劳动、服务性劳动、管理性劳动和特殊性劳动。*这五类劳动显然各有特殊性，对从事不同类型劳动的人员的知识结构、能力结构和素质结构的要求也不同。例如知识生产劳动的特点是探索性强、风险大、不准重复、周期性长。因此，对从事知识生产劳动的人，知识结构方面要求有坚实的基础理论知识和不断更新的专业知识；在能力结构方面要求有科研定向能力、发现问题的能力、较强的判断力、想象力、创造力、获得信息的能力和表达能力；在素质结构方面，要求有事业心、进取性、坚韧性、协作性和实事求是的精神。而管理性劳动则不然，由于它总是面对一个系统，并使它能有最大的输出。所以对从事管理性劳动的人员的要求，就明显地不同于从事知识生产劳动的人员。例如管理一所大学的校长，在知识结构方面，应知识面广，并能不断更新知识；在能力结构方面，应具备决策能力、组织指挥能力、综合分析能力、发现问题和解决问题的能力、知人善任的能力、协调能力、社交能力、学科发展和建设能力、人际关系能力和演讲能

---

* 特殊劳动系指教师、医生、演员、播音员之类人员所从事的劳动和部分军事活动。

力；在素质结构方面，要求有较高的政策水平、强烈的事业心和高度的责任感，有自知之明，能为人师表。

（二）不同层次、不同类别的特点

在同一种类型的劳动中，不同层次、不同类别，其特点也是不同的．为了保证人事管理考试结果客观有效，在进行考试时，应充分注意和正确解决这一问题。

同一类型的劳动，往往包括了许多不同的层次和各种不同的类别．例如知识生产劳动，有基础研究、应用研究和发展研究之分．管理性劳动，有高层次、中间层次和低层次之别．它们的每一个层次中还可以进一步细分．每一层次的劳动，由于目的不同、任务不同、责任不同，因而对人员的要求也必然不同（当然也有相同之处）．例如特殊劳动种类中的教师的劳动．虽同是教师劳动，但不同教育层次的教师劳动的目的、任务显然不一样，即使在同一教育层次也不尽相同．如小学教育，一年级教学的目的、任务就与三年级的不同，三年级与五年级也不一样．所以，为了达到人事管理考试的目的——或使人尽其才、才尽其用，或使各个岗位上的人员能更好地适应本职工作，在进行人事管理考试时，应该认真而深入地研究各种层次、各种类别劳动的特点和它对人员的要求，使考试具有针对性和促进性。

（三）不同时期的特点

人事管理考试，作为实施人事管理的重要基础，往往与晋级、革职、调资、奖惩等不同的用途有着密切的联系．不同的考试目的，应有不同的考试内容．这是人事管理考试必须注意的又一个问题。

就一般情况讲，人事管理考试可分为过去式、现在式和将来式三种。

1．过去式．通常用于晋级、调资或奖励．考试的任务是对过去的表现和绩效作出评定．在设计和确定考试的内容结构时，

重点应放在绩效方面。

2．现在式。其目的往往是对人员的晋升或提拔。这种考试往往是比较全面的，即包括素质、智能和绩效等多方面的内容。这种考试，虽然也应对被试者的过去作出评价，但考试内容应以检测现实水平为重点，并适当考虑即将就任之新工作的要求。

3．将来时。一般用于后备干部选拔和人才培养，属预测考试的范畴。这种考试的着眼点是今后一段时期内社会的需求。考试内容的设计和重点的确定，均应以下一时期社会的要求为目标。

考试作为一种社会活动具有悠久的历史。但以这种社会活动为研究对象的考试学，则是20世纪80年代的产物。它作为一门专门研究人类考试活动的科学，无论是它的学科体系还是它的研究手段和方法，都还很不完善，有待进一步的研究和探讨。况且，随着人类社会政治、经济、文化和科学技术的进一步发展，社会对人才和其他人员将会提出新的要求。为了适应和满足社会发展的需要，测度、甄选人的德、学、才、识、体个别差异的考试活动，必然会随之提出新的观点，形成新的理论，产生新的方法，出现新的手段。所以，考试学也必然有一个理论上不断丰富创新、学科体系不断发展和完善的过程。本书只是对考试学理论体系的一个初步探讨。初生之物，丑陋难免，何况一门学科体系的创立，非短时之功所能期及，而是一项长期的、浩繁艰辛的特殊劳动，常常需有一代甚至几代人的潜心探究。但是千里之行始于足下，欲引玉必先抛一砖。我们率先提出考试学理论体系的目的，正是希望广大仁人志士变丑为美，尽力缩短考试学理论体系成熟、完善的时距。

1986年12月初稿

1987年4月改定

附 录

考试常用统计数表

表1　　　T分数与百分比对照表

T分数	百分比(超过数)	T分数	百分比(超过数)
0	99.999971	60	15.87
10	99.9968	70	2.28
20	99.865	80	0.13
30	97.72	90	0.0032
40	84.13	100	0.000029
50	50.00		

表2　　百分位与标准分对照表

百分位	标准分	百分位	标准分	百分位	标准分	百分位	标准分
1	−2.328	26	−0.643	51	0.025	76	0.706
2	−2.054	27	−0.613	52	0.050	77	0.739
3	−1.881	28	−0.583	53	0.075	78	0.772
4	−1.751	29	−0.553	54	0.100	79	0.806
5	−1.645	30	−0.524	55	0.126	80	0.842
6	−1.555	31	−0.496	56	0.151	81	0.878
7	−1.476	32	−0.468	57	0.176	82	0.915
8	−1.405	33	−0.440	58	0.202	83	0.954
9	−1.341	34	−0.413	59	0.228	84	0.995
10	−1.282	35	−0.385	60	0.253	85	1.036
11	−1.227	36	−0.359	61	0.279	86	1.030
12	−1.175	37	−0.332	62	0.306	87	1.126
13	−1.126	38	−0.306	63	0.332	88	1.175
14	−1.080	39	−0.279	64	0.359	89	1.227
15	−1.036	40	−0.253	65	0.385	90	1.282
16	−0.995	41	−0.228	66	0.413	91	1.341
17	−0.954	42	−0.202	67	0.440	92	1.405
18	−0.915	43	−0.176	68	0.468	93	1.476
19	−0.878	44	−0.151	69	0.496	94	1.555
20	−0.842	45	−0.126	70	0.524	95	1.645
21	−0.806	46	−0.100	71	0.553	96	1.751
22	−0.772	47	−0.075	72	0.583	97	1.881
23	−0.739	48	−0.050	73	0.613	98	2.054
24	−0.706	49	−0.025	74	0.643	99	2.326
25	−0.675	50	0.000	75	0.675		

表3　正态曲线的面积和纵线表

标准分 (1) $z=x/\delta$	面积 (2) 自$\overline{X}$至z	积 (3) 较大部分	(4) 较小部分	纵线 (5) y
0.00	0.0000	0.5000	0.5000	0.3989
0.01	0.0040	0.5040	0.4960	0.3989
0.02	0.0080	0.5080	0.4920	0.3989
0.03	0.0120	0.5120	0.4880	0.3988
0.04	0.0160	0.5160	0.4840	0.3986
0.05	0.0199	0.5199	0.4801	0.3984
0.06	0.0239	0.5239	0.4761	0.3982
0.07	0.0279	0.5279	0.4721	0.3980
0.08	0.0319	0.5319	0.4681	0.3977
0.09	0.0359	0.5359	0.4641	0.3973
0.10	0.0398	0.5398	0.4602	0.3970
0.11	0.0438	0.5438	0.4562	0.3965
0.12	0.0478	0.5478	0.4522	0.3961
0.13	0.0517	0.5517	0.4483	0.3956
0.14	0.0557	0.5557	0.4443	0.3951
0.15	0.0596	0.5596	0.4404	0.3945
0.16	0.0636	0.5636	0.4364	0.3939
0.17	0.0675	0.5675	0.4325	0.3932
0.18	0.0714	0.5714	0.4286	0.3925
0.19	0.0753	0.5753	0.4247	0.3918
0.20	0.0793	0.5793	0.4207	0.3910
0.21	0.0832	0.5832	0.4168	0.3902
0.22	0.0871	0.5871	0.4129	0.3894
0.23	0.0910	0.5910	0.4090	0.3885
0.24	0.0948	0.5948	0.4052	0.3876
0.25	0.0987	0.5987	0.4013	0.3867
0.26	0.1026	0.6026	0.3974	0.3857
0.27	0.1064	0.6064	0.3936	0.3847
0.28	0.1103	0.6103	0.3897	0.3836
0.29	0.1141	0.6141	0.3859	0.3825
0.30	0.1179	0.6179	0.3821	0.3814
0.31	0.1217	0.6217	0.3783	0.3802
0.32	0.1255	0.6255	0.3745	0.3790
0.33	0.1293	0.6293	0.3707	0.3778
0.34	0.1331	0.6331	0.3669	0.3765

续表 3

标准分 (1) $z=x/\delta$	面积 (2) 自$\bar{X}$至z	面积 (3) 较大部分	面积 (4) 较小部分	纵线 (5) y
0.35	0.1368	0.6368	0.3632	0.3752
0.36	0.1406	0.6406	0.3594	0.3739
0.37	0.1443	0.6443	0.3557	0.3725
0.38	0.1480	0.6480	0.3520	0.3712
0.39	0.1517	0.6517	0.3483	0.3697
0.40	0.1554	0.6554	0.3446	0.3683
0.41	0.1591	0.6591	0.3409	0.3668
0.42	0.1628	0.6628	0.3372	0.3653
0.43	0.1664	0.6664	0.3336	0.3637
0.44	0.1700	0.6700	0.3300	0.3621
0.45	0.1736	0.6736	0.3264	0.3605
0.46	0.1772	0.6772	0.3228	0.3589
0.47	0.1808	0.6808	0.3192	0.3572
0.48	0.1844	0.6844	0.3156	0.3555
0.49	0.1879	0.6879	0.3121	0.3538
0.50	0.1915	0.6915	0.3085	0.3521
0.51	0.1950	0.6950	0.3050	0.3503
0.52	0.1985	0.6985	0.3015	0.3485
0.53	0.2019	0.7019	0.2981	0.3467
0.54	0.2054	0.7054	0.2946	0.3448
0.55	0.2088	0.7088	0.2912	0.3429
0.56	0.2123	0.7123	0.2877	0.3410
0.57	0.2157	0.7157	0.2843	0.3391
0.58	0.2190	0.7190	0.2810	0.3372
0.59	0.2224	0.7224	0.2776	0.3352
0.60	0.2257	0.7257	0.2743	0.3332
0.61	0.2291	0.7291	0.2709	0.3312
0.62	0.2324	0.7324	0.2676	0.3292
0.63	0.2357	0.7357	0.2643	0.3271
0.64	0.2389	0.7389	0.2611	0.3251
0.65	0.2422	0.7422	0.2578	0.3230
0.66	0.2454	0.7454	0.2546	0.3209
0.67	0.2486	0.7486	0.2514	0.3187
0.68	0.2517	0.7517	0.2483	0.3166
0.69	0.2549	0.7549	0.2451	0.3144

续表 3

标准分	面	积		纵线
(1) $z=x/\delta$	(2) 自$\bar{X}$至z	(3) 较大部分	(4) 较小部分	(5) y
0.70	0.2580	0.7580	0.2420	0.3123
0.71	0.2611	0.7611	0.2389	0.3101
0.72	0.2642	0.7642	0.2358	0.3079
0.73	0.2673	0.7673	0.2327	0.3056
0.74	0.2704	0.7704	0.2296	0.3034
0.75	0.2734	0.7734	0.2266	0.3011
0.76	0.2764	0.7764	0.2236	0.2989
0.77	0.2794	0.7794	0.2206	0.2966
0.78	0.2823	0.7823	0.2177	0.2943
0.79	0.2852	0.7852	0.2148	0.2920
0.80	0.2881	0.7881	0.2119	0.2897
0.81	0.2910	0.7910	0.2090	0.2874
0.82	0.2939	0.7939	0.2061	0.2850
0.83	0.2967	0.7967	0.2033	0.2827
0.84	0.2995	0.7995	0.2005	0.2803
0.85	0.3023	0.8023	0.1977	0.2780
0.86	0.3051	0.8051	0.1949	0.2756
0.87	0.3078	0.8078	0.1922	0.2732
0.88	0.3106	0.8106	0.1894	0.2709
0.89	0.3133	0.8133	0.1867	0.2685
0.90	0.3159	0.8159	0.1841	0.2661
0.91	0.3186	0.8186	0.1814	0.2637
0.92	0.3212	0.8212	0.1788	0.2613
0.93	0.3238	0.8238	0.1762	0.2589
0.94	0.3264	0.8264	0.1736	0.2565
0.95	0.3289	0.8289	0.1711	0.2541
0.96	0.3315	0.8315	0.1685	0.2516
0.97	0.3340	0.8340	0.1660	0.2492
0.98	0.3365	0.8365	0.1635	0.2468
0.99	0.3389	0.8389	0.1611	0.2444
1.00	0.3413	0.8413	0.1587	0.2420
1.01	0.3438	0.8438	0.1562	0.2396
1.02	0.3461	0.8461	0.1539	0.2371
1.03	0.3485	0.8485	0.1515	0.2347
1.04	0.3508	0.8508	0.1492	0.2323

续表 3

标准分	面	积		纵线
(1)	(2)	(3)	(4)	(5)
$z=x/\delta$	自$\bar{X}$至z	较大部分	较小部分	y
1.05	0.3531	0.8531	0.1469	0.2299
1.06	0.3554	0.8554	0.1446	0.2275
1.07	0.3577	0.8577	0.1423	0.2251
1.08	0.3599	0.8599	0.1401	0.2227
1.09	0.3621	0.8621	0.1379	0.2203
1.10	0.3643	0.8643	0.1357	0.2179
1.11	0.3665	0.8665	0.1335	0.2155
1.12	0.3686	0.8686	0.1314	0.2131
1.13	0.3708	0.8708	0.1292	0.2107
1.14	0.3729	0.8729	0.1271	0.2083
1.15	0.3749	0.8749	0.1251	0.2059
1.16	0.3770	0.8770	0.1230	0.2036
1.17	0.3790	0.8790	0.1210	0.2012
1.18	0.3810	0.8810	0.1190	0.1989
1.19	0.3830	0.8830	0.1170	0.1965
1.20	0.3849	0.8849	0.1151	0.1942
1.21	0.3869	0.8869	0.1131	0.1919
1.22	0.3888	0.8888	0.1112	0.1895
1.23	0.3907	0.8907	0.1093	0.1872
1.24	0.3925	0.8925	0.1075	0.1849
1.25	0.3944	0.8944	0.1056	0.1826
1.26	0.3962	0.8962	0.1038	0.1804
1.27	0.3980	0.8980	0.1020	0.1781
1.28	0.3997	0.8997	0.1003	0.1758
1.29	0.4015	0.9015	0.0985	0.1736
1.30	0.4032	0.9032	0.0968	0.1714
1.31	0.4049	0.9049	0.0951	0.1691
1.32	0.4066	0.9066	0.0934	0.1669
1.33	0.4082	0.9082	0.0918	0.1647
1.34	0.4099	0.9099	0.0901	0.1626
1.35	0.4115	0.9115	0.0885	0.1604
1.36	0.4131	0.9131	0.0869	0.1582
1.37	0.4147	0.9147	0.0853	0.1561
1.38	0.4162	0.9162	0.0838	0.1539
1.39	0.4177	0.9177	0.0823	0.1518

续表 3

标准分 (1) $z=x/\delta$	面积 (2) 自$\overline{X}$至z	面积 (3) 较大部分	面积 (4) 较小部分	纵线 (5) y
1.40	0.4192	0.9192	0.0808	0.1497
1.41	0.4207	0.9207	0.0793	0.1476
1.42	0.4222	0.9222	0.0778	0.1456
1.43	0.4236	0.9236	0.0764	0.1435
1.44	0.4251	0.9251	0.0749	0.1415
1.45	0.4265	0.9265	0.0735	0.1394
1.46	0.4279	0.9279	0.0721	0.1374
1.47	0.4292	0.9292	0.0708	0.1354
1.48	0.4306	0.9306	0.0694	0.1334
1.49	0.4319	0.9319	0.0681	0.1315
1.50	0.4332	0.9332	0.0668	0.1295
1.51	0.4345	0.9345	0.0655	0.1276
1.52	0.4357	0.9357	0.0643	0.1257
1.53	0.4370	0.9370	0.0630	0.1238
1.54	0.4382	0.9382	0.0618	0.1210
1.55	0.4394	0.9394	0.0606	0.1200
1.56	0.4406	0.9406	0.0594	0.1182
1.57	0.4418	0.9418	0.0582	0.1163
1.58	0.4429	0.9429	0.0571	0.1145
1.59	0.4441	0.9441	0.0559	0.1127
1.60	0.4452	0.9452	0.0548	0.1109
1.61	0.4463	0.9463	0.0537	0.1092
1.62	0.4474	0.9474	0.0526	0.1074
1.63	0.4484	0.9484	0.0516	0.1057
1.64	0.4495	0.9495	0.0505	0.1040
1.65	0.4505	0.9505	0.0495	0.1023
1.66	0.4515	0.9515	0.0485	0.1006
1.67	0.4525	0.9525	0.0475	0.0989
1.68	0.4535	0.9535	0.0465	0.0973
1.69	0.4545	0.9545	0.0455	0.0957
1.70	0.4554	0.9554	0.0446	0.0940
1.71	0.4564	0.9564	0.0436	0.0925
1.72	0.4573	0.9573	0.0427	0.0909
1.73	0.4582	0.9582	0.0418	0.0893
1.74	0.4591	0.9591	0.0409	0.0878

续表 3

标准分 (1) $z=x/\delta$	面积 (2) 自$\overline{X}$至z	(3) 较大部分	(4) 较小部分	纵线 (5) y
1.75	0.4599	0.9599	0.0401	0.0863
1.76	0.4608	0.9608	0.0392	0.0848
1.77	0.4616	0.9616	0.0384	0.0833
1.78	0.4625	0.9625	0.0375	0.0818
1.79	0.4633	0.9633	0.0367	0.0804
1.80	0.4641	0.9641	0.0359	0.0790
1.81	0.4649	0.9649	0.0351	0.0775
1.82	0.4656	0.9656	0.0344	0.0761
1.83	0.4664	0.9664	0.0336	0.0748
1.84	0.4671	0.9671	0.0329	0.0734
1.85	0.4678	0.9678	0.0322	0.0721
1.86	0.4686	0.9686	0.0314	0.0707
1.87	0.4693	0.9693	0.0307	0.0694
1.88	0.4699	0.9699	0.0301	0.0681
1.89	0.4706	0.9706	0.0294	0.0669
1.90	0.4713	0.9713	0.0287	0.0656
1.91	0.4719	0.9719	0.0281	0.0644
1.92	0.4726	0.9726	0.0274	0.0632
1.93	0.4732	0.9732	0.0268	0.0620
1.94	0.4738	0.9738	0.0262	0.0608
1.95	0.4744	0.9744	0.0256	0.0596
1.96	0.4750	0.4750	0.0250	0.0584
1.97	0.4756	0.9756	0.0244	0.0573
1.98	0.4761	0.9761	0.0239	0.0562
1.99	0.4767	0.9767	0.0233	0.0551
2.00	0.4772	0.9772	0.0228	0.0540
2.01	0.4778	0.9778	0.0222	0.0529
2.02	0.4783	0.9783	0.0217	0.0519
2.03	0.4788	0.9788	0.0212	0.0508
2.04	0.4793	0.9793	0.0207	0.0498
2.05	0.4798	0.9798	0.0202	0.0488
2.06	0.4803	0.9803	0.0197	0.0478
2.07	0.4808	0.9808	0.0192	0.0468
2.08	0.4812	0.9812	0.0188	0.0459
2.09	0.4817	0.9817	0.0183	0.0449

续表 3

标准分数 (1) $z=x/\delta$	面积 (2) 自 $\bar{X}$ 至 z	(3) 较大部分	(4) 较小部分	纵线 (5) y
2.10	0.4821	0.9821	0.0179	0.0440
2.11	0.4826	0.9826	0.0174	0.0431
2.12	0.4830	0.9830	0.0170	0.0422
2.13	0.4834	0.9834	0.0166	0.0413
2.14	0.4838	0.9838	0.0162	0.0404
2.15	0.4842	0.9842	0.0158	0.0396
2.16	0.4846	0.9846	0.0154	0.0387
2.17	0.4850	0.9850	0.0150	0.0379
2.18	0.4854	0.9854	0.0146	0.0371
2.19	0.4857	0.9857	0.0143	0.0363
2.20	0.4861	0.9861	0.0139	0.0355
2.21	0.4864	0.9864	0.0136	0.0347
2.22	0.4868	0.9868	0.0132	0.0339
2.23	0.4871	0.9871	0.0129	0.0332
2.24	0.4875	0.9875	0.0125	0.0325
2.25	0.4878	0.9878	0.0122	0.0317
2.26	0.4881	0.9881	0.0119	0.0310
2.27	0.4884	0.9884	0.0116	0.0303
2.28	0.4887	0.9887	0.0113	0.0297
2.29	0.4890	0.9890	0.0110	0.0290
2.30	0.4893	0.9893	0.0107	0.0283
2.31	0.4896	0.9896	0.0104	0.0277
2.32	0.4898	0.9898	0.0102	0.0270
2.33	0.4901	0.9901	0.0099	0.0264
2.34	0.4904	0.9904	0.0096	0.0258
2.35	0.4906	0.9906	0.0094	0.0252
2.36	0.4909	0.9909	0.0091	0.0246
2.37	0.4911	0.9911	0.0089	0.0241
2.38	0.4913	0.9913	0.0087	0.0235
2.39	0.4916	0.9916	0.0084	0.0229
2.40	0.4918	0.9918	0.0082	0.0224
2.41	0.4920	0.9920	0.0080	0.0219
2.42	0.4922	0.9922	0.0078	0.0213
2.43	0.4925	0.9925	0.0075	0.0208
2.44	0.4927	0.9927	0.0073	0.0203

续表 3

标准分 (1) $z=x/\delta$	面积 (2) 自$\overline{X}$至z	面积 (3) 较大部分	面积 (4) 较小部分	纵线 (5) y
2.45	0.4929	0.9929	0.0071	[illegible]
2.46	0.4931	0.9931	0.0069	0.0194
2.47	0.4932	0.9932	0.0068	0.0189
2.48	0.4934	0.9934	0.0066	0.0184
2.49	0.4936	0.9936	0.0064	0.0180
2.50	0.4938	0.9938	0.0062	0.0175
2.51	0.4940	0.9940	0.0060	0.0171
2.52	0.4941	0.9941	0.0059	0.0167
2.53	0.4943	0.9943	0.0057	0.0163
2.54	0.4945	0.9945	0.0055	0.0158
2.55	0.4946	0.9946	0.0054	0.0154
2.56	0.4948	0.9948	0.0052	0.0151
2.57	0.4949	0.9949	0.0051	0.0147
2.58	0.4951	0.9951	0.0049	0.0143
2.59	0.4952	0.9952	0.0048	0.0139
2.60	0.4953	0.9953	0.0047	0.0136
2.61	0.4955	0.9955	0.0045	0.0132
2.62	0.4956	0.9956	0.0044	0.0129
2.63	0.4957	0.9957	0.0043	0.0126
2.64	0.4959	0.9959	0.0041	0.0122
2.65	0.4960	0.9960	0.0040	0.0119
2.66	0.4961	0.9961	0.0039	0.0116
2.67	0.4962	0.9962	0.0038	0.0113
2.68	0.4963	0.9963	0.0037	0.0110
2.69	0.4964	0.9964	0.0036	0.0107
2.70	0.4965	0.9965	0.0035	0.0104
2.71	0.4966	0.9966	0.0034	0.0101
2.72	0.4967	0.9967	0.0033	0.0099
2.73	0.4968	0.9968	0.0032	0.0096
2.74	0.4969	0.9969	0.0031	0.0093
2.75	0.4970	0.9970	0.0030	0.0091
2.76	0.4971	0.9971	0.0029	0.0088
2.77	0.4972	0.9972	0.0028	0.0086
2.78	0.4973	0.9973	0.0027	0.0084
2.79	0.4974	0.9974	0.0026	0.0081

续表 3

标准分 (1) $z=x/\delta$	面积 (2) 自$\bar{X}$至z	面积 (3) 较大部分	面积 (4) 较小部分	纵线 (5) y
2.80	0.4974	0.9974	0.0026	0.0079
2.81	0.4975	0.9975	0.0025	0.0077
2.82	0.4976	0.9976	0.0024	0.0075
2.83	0.4977	0.9977	0.0023	0.0073
2.84	0.4977	0.9977	0.0023	0.0071
2.85	0.4978	0.9978	0.0022	0.0069
2.86	0.4979	0.9979	0.0021	0.0067
2.87	0.4979	0.9979	0.0021	0.0065
2.88	0.4980	0.9980	0.0020	0.0063
2.89	0.4981	0.9981	0.0019	0.0061
2.90	0.4981	0.9981	0.0019	0.0060
2.91	0.4982	0.9982	0.0018	0.0058
2.92	0.4982	0.9982	0.0018	0.0056
2.93	0.4983	0.9983	0.0017	0.0055
2.94	0.4984	0.9984	0.0016	0.0053
2.95	0.4984	0.9984	0.0016	0.0051
2.96	0.4985	0.9985	0.0015	0.0050
2.97	0.4985	0.9985	0.0015	0.0048
2.98	0.4986	0.9986	0.0014	0.0047
2.99	0.4986	0.9986	0.0014	0.0046
3.00	0.4987	0.9987	0.0013	0.0044
3.01	0.4987	0.9987	0.0013	0.0043
3.02	0.4987	0.9987	0.0013	0.0042
3.03	0.4988	0.9988	0.0012	0.0040
3.04	0.4988	0.9988	0.0012	0.0039
3.05	0.4989	0.9989	0.0011	0.0038
3.06	0.4989	0.9989	0.0011	0.0037
3.07	0.4989	0.9989	0.0011	0.0036
3.08	0.4990	0.9990	0.0010	0.0035
3.09	0.4990	0.9990	0.0010	0.0034
3.10	0.4990	0.9990	0.0010	0.0033
3.11	0.4991	0.9991	0.0009	0.0032
3.12	0.4991	0.9991	0.0009	0.0031
3.13	0.4991	0.9991	0.0009	0.0030
3.14	0.4992	0.9992	0.0008	0.0029

续表 3

标准分	面	积		纵线
(1)	(2)	(3)	(4)	(5)
$z=x/\delta$	自$\bar{X}$至z	较大部分	较小部分	y
3.15	0.4992	0.9992	0.0008	0.0028
3.16	0.4992	0.9992	0.0008	0.0027
3.17	0.4992	0.9992	0.0008	0.0026
3.18	0.4993	0.9993	0.0007	0.0025
3.19	0.4993	0.9993	0.0007	0.0025
3.20	0.4993	0.9993	0.0007	0.0024
3.21	0.4993	0.9993	0.0007	0.0023
3.22	0.4994	0.9994	0.0006	0.0022
3.23	0.4994	0.9994	0.0006	0.0022
3.24	0.4994	0.9994	0.0006	0.0021
3.30	0.4995	0.9995	0.0005	0.0017
3.40	0.4997	0.9997	0.0003	0.0012
3.50	0.4998	0.9998	0.0002	0.0009
3.60	0.4998	0.9998	0.0002	0.0006
3.70	0.4999	0.9999	0.0001	0.0004

表4　　r 值的Z转换表

r	Zr		Zr	r	Zr	r	Zr	r	Zr
.000	.000	.200	.203	.400	.424	.600	.693	.800	1.099
.005	.005	.205	.208	.405	.430	.605	.701	.805	1.113
.010	.010	.210	.213	.410	.436	.610	.709	.810	1.127
.015	.015	.215	.218	.515	.442	.615	.717	.815	1.142
.020	.020	.220	.224	.420	.448	.620	.725	.820	1.157
.025	.025	.225	.229	.425	.454	.625	.733	.825	1.172
.030	.030	.230	.234	.430	.460	.630	.741	.830	1.188
.035	.035	.235	.239	.435	.466	.635	.750	.835	1.204
.040	.040	.240	.245	.440	.472	.640	.758	.840	1.221
.045	.045	.245	.250	.445	.478	.645	.767	.845	1.238
.050	.050	.250	.255	.450	.485	.650	.775	.850	1.256
.055	.055	.255	.261	.455	.491	.655	.784	.855	1.274
.060	.060	.260	.266	.460	.497	.660	.793	.860	1.253
.065	.065	.265	.271	.465	.504	.665	.802	.865	1.313
.070	.070	.270	.277	.470	.510	.670	.811	.870	1.33[illegible]
.075	.075	.275	.282	.475	.517	.675	.820	.875	1.354
.080	.080	.280	.288	.480	.523	.680	.829	.880	1.376
.085	.085	.285	.293	.485	.530	.685	.838	.885	1.398
.090	.090	.290	.299	.490	.536	.690	.848	.890	1.422
.095	.095	.295	.304	.495	.543	.695	.858	.895	1.447
.100	.100	.300	.310	.500	.549	.700	.867	.900	1.472
.105	.105	.305	.315	.505	.556	.705	.877	.905	1.499
.110	.110	.310	.321	.510	.563	.710	.887	.910	1.528
.115	.116	.315	.326	.515	.570	.715	.897	.915	1.557
.120	.121	.320	.332	.520	.576	.720	.908	.920	1.589
.125	.126	.325	.337	.525	.583	.725	.918	.925	1.623
.130	.131	.330	.343	.530	.590	.730	.929	.930	1.658
.135	.136	.335	.348	.535	.597	.735	.940	.935	1.697
.140	.141	.340	.354	.540	.604	.740	.950	.940	1.738
.145	.146	.345	.360	.545	.611	.745	.962	.945	1.783
.150	.151	.350	.365	.550	.618	.750	.973	.950	1.832
.155	.156	.355	.371	.555	.626	.755	.984	.955	1.886
.160	.161	.360	.377	.560	.633	.760	.996	.960	1.946
.165	.167	.365	.383	.565	.640	.765	1.008	.965	2.014
.170	.172	.370	.388	.570	.648	.770	1.020	.970	2.092
.175	.177	.375	.394	.575	.655	.775	1.033	.975	2.185
.180	.182	.380	.400	.580	.662	.780	1.045	.980	2.298
.185	.187	.385	.406	.585	.670	.785	1.058	.985	2.443
.190	.192	.390	.412	.590	.678	.790	1.071	.990	2.647
.195	.198	.395	.418	.595	.685	.795	1.085	.995	2.994

表 5　　等级相关系数界值表

n	P(2):0.50 P(1):0.25	0.20 0.10	0.10 0.05	0.05 0.025	0.02 0.01	0.01 0.005	0.005 0.0025	0.002 0.001	0.001 0.0005
4	0.600	1.000	1.000						
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.217	0.406	0.503	0.587	0.678	0.727	0.769	0.818	0.846
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.791	0.824
14	0.200	0.367	0.464	0.538	0.626	0.679	0.723	0.771	0.802
15	0.189	0.354	0.446	0.521	0.604	0.654	0.700	0.750	0.779
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.729	0.762
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.3[illegible]9	0.391	0.460	0.535	0.[illegible]84	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.662	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.514	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.431	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580
31	0.126	0.236	0.301	0.356	0.418	0.459	0.496	0.541	0.571
32	0.124	0.232	0.296	0.350	0.412	0.452	0.489	0.533	0.563
33	0.121	0.229	0.291	0.345	0.405	0.446	0.482	0.525	0.554
34	0.120	0.225	0.287	0.340	0.399	0.439	0.475	0.517	0.547
35	0.118	0.222	0.283	0.335	0.394	0.433	0.468	0.510	0.539
36	0.116	0.219	0.279	0.330	0.388	0.427	0.462	0.504	0.533
37	0.114	0.216	0.275	0.325	0.383	0.421	0.456	0.497	0.526
38	0.113	0.212	0.271	0.321	0.378	0.415	0.450	0.491	0.519
39	0.111	0.210	0.267	0.317	0.373	0.410	0.444	0.485	0.513
40	0.110	0.207	0.264	0.313	0.368	0.405	0.439	0.479	0.507
41	0.108	0.204	0.261	0.309	0.364	0.400	0.433	0.473	0.501
42	0.107	0.202	0.257	0.305	0.359	0.395	0.428	0.468	0.495
43	0.105	0.190	0.254	0.301	0.355	0.391	0.423	0.463	0.490
44	0.104	0.197	0.251	0.298	0.351	0.386	0.419	0.458	0.484
45	0.103	0.194	0.248	0.294	0.347	0.382	0.414	0.453	0.479
46	0.102	0.192	0.246	0.291	0.343	0.378	0.410	0.448	0.474
47	0.101	0.190	0.243	0.288	0.340	0.374	0.405	0.443	0.469
48	0.100	0.188	0.240	0.285	0.336	0.370	0.401	0.439	0.465
49	0.098	0.186	0.238	0.282	0.333	0.366	0.397	0.434	0.460
50	0.097	0.184	0.235	0.279	0.329	0.363	0.393	[illegible]430	0.456

表6　　积差相关系数显著临界值表

$df=N-2$	$\alpha=.10$	.05	.02	.01
1	.988	.997	.9995	.9999
2	.900	.950	.980	.990
3	.805	.878	.934	.959
4	.729	.811	.882	.917
5	.669	.754	.833	.874
6	.622	.707	.789	.834
7	.582	.666	.750	.798
8	.549	.632	.716	.765
9	.521	.602	.685	.735
10	.497	.576	.658	.708
11	.476	.553	.634	.684
12	.458	.532	.612	.661
13	.441	.514	.592	.641
14	.426	.497	.574	.623
15	.412	.482	.558	.606
16	.400	.468	.542	.590
17	.389	.456	.528	.575
18	.378	.444	.516	.561
19	.369	.433	.503	.549
20	.360	.423	.492	.537
21	.352	.413	.482	.526
22	.341	.404	.472	.515
23	.337	.396	.462	.505
24	.330	.388	.453	.496
25	.323	.381	.445	.487
26	.317	.374	.437	.479
27	.311	.367	.430	.471
28	.306	.361	.423	.463
29	.301	.355	.416	.456
30	.296	.349	.409	.449
35	.275	.325	.381	.418
40	.257	.304	.358	.393
45	.243	.288	.338	.372
50	.231	.273	.322	.354
60	.211	.250	.295	.325
70	.195	.232	.274	.302
80	.183	.217	.256	.283
90	.173	.205	.242	.267
100	.164	.195	.230	.254

表 7 弗拉南根二列相关系数估计数表

下端27%答对比率 \ 上端27%答对比率	02	06	10	14	18	22	26	30	34	38	42	46	50	54	58	62	66	70	74	78	82	86	90	94	98
02	00	19	30	37	43	48	51	55	58	61	63	66	68	70	72	73	75	77	79	80	82	84	86	88	98
06		00	11	19	26	31	36	40	44	47	50	53	56	59	61	64	66	68	71	73	76	78	81	84	88
10			00	08	15	21	26	30	34	38	41	45	48	51	54	57	60	63	65	68	71	74	77	81	86
14				00	07	12	18	22	27	31	34	38	42	45	48	51	54	57	60	63	67	70	74	78	84
18					00	06	11	16	20	25	28	32	36	39	43	47	49	53	56	60	63	67	71	76	82
22						00	06	10	15	19	23	27	31	34	38	42	45	49	52	56	60	63	66	73	80
26							00	05	09	14	18	22	26	30	33	37	41	44	48	52	56	60	65	71	79
30								00	04	09	13	17	21	25	29	33	37	40	44	49	53	57	63	68	77
34									00	04	09	13	17	21	25	29	33	37	41	45	49	54	60	66	75
38										00	04	08	13	16	20	25	29	33	37	42	47	51	57	64	73
42											00	04	08	12	16	20	25	29	33	38	43	48	54	61	72
46												00	04	08	12	16	20	25	30	34	39	45	51	59	70
50													00	04	08	13	16	21	26	31	36	42	48	56	68
54														00	04	08	13	17	22	27	32	38	45	53	66
58															00	04	08	13	18	23	28	34	41	50	63
62																00	04	09	14	19	25	31	38	47	61
66																	00	04	09	15	20	27	34	44	58
70																		00	05	10	16	22	30	40	55
74																			00	06	11	18	26	30	51
78																				00	06	12	21	31	48
82																					00	07	15	26	43
86																						00	08	19	37
90																							00	11	30
94																								00	19
98																									00

主要参考文献

《马克思恩格斯全集》第3、23、26卷，人民出版社1972年。
《马克思恩格斯选集》第1、2、3、4卷，人民出版社1972年。
《列宁全集》第2、5、23、35卷，人民出版社1959年。
《列宁选集》第2、4卷，人民出版社1975年。
马克思：《资本论》第1卷，人民出版社1975年。
斯大林：《辩证唯物主义与历史唯物主义》，人民出版社1953年。
《毛泽东选集》（合订本），人民出版社1971年。
李 达：《唯物辩证法大纲》，人民出版社1978年。
郑玄注：《周礼》。
司马迁：《史记》。
班 固：《汉书》。
杜 佑：《通典》。
郑 樵：《通志》。
马端临：《文献通考》。
范 晔：《后汉书》。
陈 寿：《三国志》。
司马光：《资治通鉴》。
翦伯赞：《中国史纲要》，人民出版社1983年。
雷祯孝：《中国人才思想史》，中国展望出版社1986年。
林举岱：《世界近代史》，上海人民出版社1984年。
张晋藩等：《科举制度史话》，中华书局1980年。
崔连仲：《世界史》（古代史），人民出版社1983年。
考试院考铨丛书指导委员会主编：《中华民国考选制度》，（台）正中书局1983年。
〔美〕亚尔保德·兰著，浦漪人等译：《新法考试》，正中书局1935年。
杨百揆等：《西方文官系统》，四川人民出版社1985年。

金观涛等：《西方社会结构的演变》，四川人民出版社1985年。
〔美〕阿历克斯·英格尔斯等著，殷陆君译：《人的现代化》，四川人民出版社1985年。
〔德〕恩斯特·卡西尔著，甘阳译：《人论》，上海译文出版社版1985年。
〔英〕罗伯特·蒙哥马利著，黄鸣译：《考试的新探索》，广西人民出社1984年。
〔奥〕弗洛伊德著，孙恺祥译：《论创造力与无意识》，中国展望出版社1986年。
罗国杰：《马克思主义伦理学》，人民出版社1986年。
〔苏〕A·H·鲁克著，周义澄等译：《创造心理学概论》，黑龙江人民出版社1985年。
〔美〕赖特·米尔斯等著，何维凌等译：《社会学与社会组织》，浙江人民出版社1986年。
〔南〕斯·布里舍里奇著，贾春增等译：《社会学原理》，东方出版社出版1986年。
吴 铎：《简明社会学》，华东师范大学出版社1986年。
时蓉华：《社会心理学》，上海人民出版社1986年。
华中师范大学教育系等：《德育学》，陕西人民出版社1986年。
李晓明：《模糊性：人类认识之谜》，人民出版社1985年。
王通讯：《人才学通论》，天津人民出版社1986年。
朱钧侃：《现代人才管理学》，南京工学院出版社1986年。
欧炳进等：《人事管理学》，广东高等教育出版社1986年。
王 康等：《人才知识手册》，湖北科技出版社1986年。
国务院办公厅调查研究室编：《中国行政管理学初探》，经济出版社1984年。
莫吉武：《社会主义领导学》，华中师范大学出版社1986年。
孟宪承等：《中国古代教育史资料》，人民出版社1985年。
〔美〕E·G·波林著，高觉敷译：《实验心理学》，商务印书馆1981年。
孙邦正：《教育测验及统计》，(台)商务印书馆1986年。

李聪明：《教育评价的理论与方法》，(台)幼狮文化事业出版公司1983年。
郭文新：《教育统计法》，（台）正中书局1977年。
〔美〕H·L·阿尔德等：《概率与统计导论》，北京大学出版社1984年。
清华大学自然辩证法教研组编：《科学技术史讲义》，清华大学出版社1984年。
潘懋元：《高等教育学》，人民教育出版社、福建教育出版社1984年。
〔苏〕季亚琴科等著，徐世京等译：《高等学校心理学》，上海译文出版社1985年。
〔美〕杜·舒尔茨著，彭聃龄等译：《现代心理学史》，人民出版社1981年。
高觉敷：《西方近代心理学史》，人民出版社1982年。
王孝玲：《教育统计学》，华东师范大学出版社1986年。
潘 菽：《教育心理学》，人民教育出版社1980年。
朱智贤等：《思惟发展心理学》，北京师范大学出版社1986年。
〔苏〕格烈米亚茨基著，王平等译：《人体解剖学》，高等教育出版社1956年。
中国医科大学主编：《人体解剖学》，人民卫生出版社1978年。
山东师范学院等合编：《生理卫生学》，人民教育出版社1977年。
〔美〕R·F·汤普森主编，孙晔等编译：《生理心理学》，科学出版社1981年。
朱智贤：《儿童心理学》，人民教育出版社1981年。
〔日〕山内光哉编著，李蔚楚等译：《学习与教学心理学》，教育科学出版社1986年。
〔美〕卡扎米亚斯等著，福建师大教育系等合译：《教育的传统与变革》，文化出版社1981年。
邢文华等著：《中小学生体质测定与评价》，人民出版社1983年。
王承绪等主编：《比较教育》，人民教育出版社1986年。
朱有瓛主编：《中国近代学制史料》，华东师范大学出版社1983年。
〔英〕J·D·尼斯比特等著，张渭诚等译：《教育研究法》，教育科学

出版社1982年。
华东师大教育学院等编：《新技术革命与教育》，华东师范大学出版社1984年。
中国科学院数学研究所统计组编，《常用数理统计方法》，科学出版社1973年。
洪 帆：《离散数学基础》，华中工学院出版社1984年。
李庆扬等：《数值分析》，华中工学院出版社1983年。
《经营管理知识丛书》之九，《智力开发》，上海交通大学出版社1986年。
黄顺基等主编：《大杠杆》，山东大学出版社1985年。
夏禹龙等：《现代智囊团》，知识出版社1984年。
〔苏〕哥拉巴里等著，王孝玲译：《数理统计在教育研究中的应用》，华东师范大学出版社1986年。
陈选善：《教育测验》，商务印书馆1935年。